

# Random friend trees

Louigi Addario-Berry\*      Simon Briend†      Luc Devroye‡  
Serte Donderwinkel §      Céline Kerriou ¶      Gábor Lugosi ||

## Abstract

We study a random recursive tree model featuring complete redirection called the random friend tree and introduced by Saramäki and Kaski. Vertices are attached in a sequential manner one by one by selecting an existing target vertex and connecting to one of its neighbours (or friends), chosen uniformly at random. This model has interesting emergent properties, such as a highly skewed degree sequence. In contrast to the preferential attachment model, these emergent phenomena stem from a local rather than a global attachment mechanism. The structure of the resulting tree is also strikingly different from both the preferential attachment tree and the uniform random recursive tree: every edge is incident to a macro-hub of asymptotically linear degree, and with high probability all but at most  $n^{9/10}$  vertices in a tree of size  $n$  are leaves. We prove various results on the neighbourhood of fixed vertices and edges, and we study macroscopic properties such as the diameter and the degree distribution, providing insights into the overall structure of the tree. We also present a number of open questions on this model and related models.

**Keywords:** Random friend trees; random recursive trees; random trees; condensation.

**MSC2020 subject classifications:** 60C05;60J80;05C05.

---

\*Department of Mathematics and Statistics, McGill University, Montréal, Québec, Canada. E-mail: [louigi.addario@mcgill.ca](mailto:louigi.addario@mcgill.ca)

†Université Paris-Saclay, CNRS, Laboratoire de Mathématiques d'Orsay, 91404, Orsay, France. E-mail: [simon.briend@universite-paris-saclay.fr](mailto:simon.briend@universite-paris-saclay.fr)

‡School of Computer Science, McGill University, Montréal, Québec, Canada. E-mail: [lucdevroye@gmail.com](mailto:lucdevroye@gmail.com)

§Bernoulli Institute and CogniGron (Groningen Cognitive Systems and Materials Center), University of Groningen (Univ Groningen), Nijenborgh 4, NL-9747 AG Groningen, Netherlands. E-mail: [s.a.donderwinkel@rug.nl](mailto:s.a.donderwinkel@rug.nl)

¶Department of Mathematics and Computer Science, Universität zu Köln, Weyertal 86-90, 50931 Cologne, Germany. E-mail: [ckerriou@math.uni-koeln.de](mailto:ckerriou@math.uni-koeln.de)

||Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain; ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain; Barcelona Graduate School of Economics. E-mail: [gabor.lugosi@gmail.com](mailto:gabor.lugosi@gmail.com)

## 1 Introduction

**Growing networks.** Various real-life phenomena, including contagion, social networks, rumour spreading, and the internet, have been described by models of growing networks (see e.g. Kumar et al. [16]). Among these models, preferential attachment, introduced by Barabási and Albert [3] is arguably the most well-studied. In this model, vertices arrive one by one, and at each time a new vertex connects to one or more existing vertices with probability proportional to their degree. The degree sequence of this model satisfies the so-called scale-free property, which is often also observed in real-world models. In contrast to models such as the configuration model or the inhomogeneous random graph model, this property of the degree sequence is an intrinsic result of the dynamics rather than a pre-imposed characteristic. This makes the preferential attachment and its related models popular tools for understanding why real-world models may develop in this way. These models however require the knowledge of the full degree sequence in order to attach a new vertex. This requirement is unnatural for real-world networks and impractical in implementation.

**The model.** The friend tree is a randomly growing network of which the dynamic also autonomously produces highly skewed degree sequences, but whose attachment rule is based on redirection and requires only local information. Models involving redirection were introduced by Kleinberg et al. [14] in directed, rooted graphs. In these models, a new vertex connects to a uniformly random vertex, or, with probability  $p$ , it connects to the ancestor of a randomly selected vertex. This mechanism, called *directed redirection*, yields a shifted linear preferential attachment rule, where the new vertex connects to a vertex with degree  $d$  with probability proportional to  $d - 2 + 1/p$ . A variant was studied by Banerjee et al. [2], where a new vertex attaches to the graph by randomly sampling a vertex and attaching to the endpoint of a path of random length directed towards the root. The undirected version of the model that we study was introduced by Saramäki and Kaski [19] and yields strikingly different graphs. In the works of Saramäki and Kaski [19] and Evans and Saramäki [10], the authors make the claim that the tree has the same law as a preferential attachment tree. This turns out to be inaccurate, as was noted by Cannings and Jordan [5]. In the undirected version, newly added vertices connect to a neighbour of a randomly selected vertex. More precisely, the starting tree  $T_2$  consists of a single edge joining vertices labelled 1 and 2. Inductively, for  $n \geq 2$ , let  $V_n \in \{1, \dots, n\}$  be chosen uniformly at random and let  $W_n$  be a uniformly random neighbour of  $V_n$  in  $T_n$ . Then build  $T_{n+1}$  from  $T_n$  by attaching a new vertex labelled  $n + 1$  to the vertex  $W_n$ , see Figure 1. Note that, for all  $n \geq 2$ , the vertex set of  $T_n$  is  $\{1, \dots, n\}$ . Moreover, setting  $W_1 = 1$ , then the edge set of  $T_n$  is  $\{\{m + 1, W_m\}, 1 \leq m \leq n - 1\}$ . We call  $T_n$  the *1-step friend tree*, inspired by the following picture. In a 1-step friend tree a person selects a random stranger and befriends a uniformly random friend of theirs. A 2-step friend tree would correspond to, instead, befriending a uniformly random friend of the stranger's random friend. In most of this work, we refer to the *1-step friend tree* simply as the *random friend tree (RFT)*.

**Motivation and challenges.** This model gives rise to numerous interesting emergent phenomena that make it worth studying. A first feature of the model is a rich-get-richer mechanism. In the preferential attachment model, it is embedded into the dynamics that vertices with many neighbours accumulate more neighbours. In the friend tree, vertices that have many neighbours *with low degree* accumulate more neighbours. It turns out that, the highest degree vertices mostly have leaves as neighbours, so the growth of high-degree vertices is mostly determined by their degree. In fact, the reinforcement is strengthened by a second effect: the larger the degree of a vertex, the less likely it is for its neighbours to increase their degree, so the more likely it is for the high degree

## Random friend trees

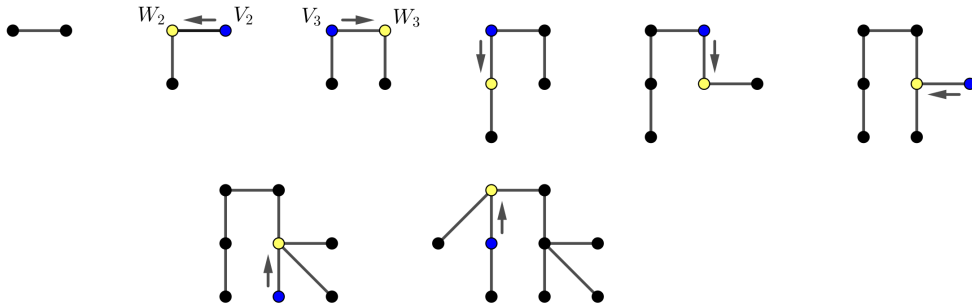


Figure 1: A realisation up to  $n = 9$ , with  $V_n$  in blue and  $W_n$  in yellow.

vertex itself to increase its degree. Unlike in the preferential attachment model, this is a result of the dynamics rather than a built-in feature.

Furthermore, the model is part of a whole family of models that in some sense interpolate between the uniform random recursive tree (URRT) and the linear preferential attachment tree (PA tree). Indeed, in the  $k$ -step friend tree, new vertices attach to the endpoint of a random walk of length  $k$  that starts at a random vertex. If  $k = 0$ , the resulting model is the URRT. If  $k$  was chosen so large that the random walk is perfectly mixed, the resulting model would be a PA tree. In other words, for  $k$  large, the starting point of the random walk has a negligible effect and the end point is distributed proportionally to the degree of each vertex. Our work demonstrates several features of the 1-step friend tree which are remarkably different from both the URRT and the PA tree. It is therefore worth investigating what range of behaviour can be observed in the entire family. In a related model, Engländer et al. [9] study a “random walk tree builder” where a tree is grown by a random walk. More precisely, a walker is moving at random on the tree and at each time step  $n$ , with probability  $n^{-\gamma}$ , a neighbour is added to the vertex where the walker is. They prove that this model is actually a PA tree for an appropriate choice of  $\gamma$ . Unlike the friend tree model, where a new random walk is started at every time step, a single random walk is able to produce a tree displaying a rich gets richer phenomenon. This further motivates the study of the friend tree model in its whole range, and suggests the investigation of its possible links to the “random walk tree builder” model.

The challenges of studying 1-step friend trees are numerous. For example, even if the process grows locally (one needs to know the neighbours of the randomly picked vertex to understand the connection probabilities), tracking only local information does not suffice to study how degrees evolve over time. Indeed, to understand how the vertex degrees change within two time steps, one must keep track of the degrees of second neighbours, and in general, for  $t$  time steps one must track the  $t$ 'th neighbourhood of every vertex. Note that this challenge does not arise in either directed redirection or the preferential attachment model; in those models the growth of a vertex degree only depends on the vertex degree itself, so degrees can be tracked on their own without considering the global structure of the tree. For  $k$ -step friend trees, the dependencies grow stronger as  $k$  increases, bringing in new challenges that we do not attempt to tackle in the current work.

**Results and comparison.** Some of the structural properties of random friend trees are comparable to those of the URRT. In Theorem 4.2, we show that the diameter is of logarithmic order almost surely (like for the URRT and the PA tree [18]). Moreover, as we show in Theorems 4.3 and 4.4 respectively, both in the random friend tree and in the

URRT, the largest distance to the nearest leaf in the  $n$ -vertex tree is  $\Theta(\log(n)/\log \log(n))$  in probability.

However, the interaction between neighbouring vertices in the attachment procedure yields significant structural differences between the random friend tree and both the URRT and the PA tree. The degree sequence might be the most illustrative of this difference. Regarding the high-degree vertices, we prove in Theorem 3.2 that “hubs” of linear degree appear almost surely, whereas in a URRT the maximum degree is logarithmic (Devroye and Lu [7]) and in linear preferential attachment tree the largest degree is of order  $\sqrt{n}$  (van der Hofstad [20], Theorem 1.17). In fact, the dynamics of low-degree vertices ‘feeding their neighbours’ implies that for every edge, at least one of the endpoints has asymptotically linear degree almost surely, so that a highly modular network emerges. This phenomenon has not been observed in any other random tree model, as far as the authors of this paper are aware. The existence of linear degree hubs also implies that two uniformly random vertices in  $T_n$  are at distance two from each other with probability bounded away from zero, while in both the URRT and the PA tree typical distances grow logarithmically [6],[21, Theorem 8.1]. As for low-degree vertices, both in PA trees and URRT an asymptotically positive fraction of vertices have degree at least two, but we shall show that a random friend tree of size  $n$  has  $n - o(n^{0.9})$  leaves. While a URRT of size  $n$  has on average  $n/2^{k-1}$  vertices of degree at least  $k$  for fixed  $k$  (Janson [11]), for friend trees, asymptotically, this number sits between  $n^{0.1}$  and  $n^{0.9}$  (see Theorem 4.7). The proliferation of hubs and the interaction between neighbours block most leaves from ever growing their degree. Proposition 5.1 shows that most leaves remain leaves forever<sup>1</sup>, whereas for URRT and PA trees, the degree of every vertex is a.s. unbounded.

**Earlier work.** The only previous rigorous result on random friend trees the authors are aware of was obtained by Cannings and Jordan [5], who show that in the 1-step friend tree,  $n - o(n)$  of the vertices are leaves almost surely. Random friend trees were also studied in the physics literature by Krapivsky and Redner [15]. In that work, the authors use simulations and non-rigorous arguments to study the distribution of size of the largest degrees and the order of growth of the number of non-leaves, and estimate the degree distribution restricted to the bounded-degree vertices. They conjecture that, for any fixed  $k$ , the number of degree  $k$  vertices is of the order of  $n^\mu$  for  $\mu \approx 0.566$ . Moreover, they conjecture that among the non-leaf vertices, the proportion of degree  $k$  vertices is of the order of  $k^{-(1+\mu)}$ . We discuss their estimates in Section 7.

**Outline.** First, in Section 2 we introduce notation that we use throughout the paper. We then present our main results for friend trees. Our findings can naturally be divided into local and global properties, which we present in Section 3 and 4 respectively. In Sections 5 and 6 we prove our main results. Finally, Section 7 contains some open questions about random friend trees.

## 2 Notation

For a graph  $G$  and a vertex  $v$  of  $G$ , write  $\mathcal{N}(v; G)$  for the neighbourhood of  $v$  in  $G$  and  $\mathcal{L}(v; G)$  for the set of leaf neighbours of  $v$  in  $G$  (i.e. vertices of degree one in  $\mathcal{N}(v; G)$ ). In the rest of the paper,  $T_n$  denotes a tree of size  $n$  obtained from the random friend tree model. The set of vertices of  $T_n$  is  $[n] := \{1, \dots, n\}$ , where the label of the vertex is its time of arrival in the tree. Since every integer  $k \in \mathbb{N}$  is a vertex of  $T_n$  for all  $n \geq k$ , we take the liberty of referring to integers of  $\mathbb{N}$  as vertices. For  $v \in T_n$ , let  $D_n(v) = |\mathcal{N}(v; T_n)|$  be the degree and let  $L_n(v) = |\mathcal{L}(v; T_n)|$  be the number of leaf

---

<sup>1</sup>So one might argue that ‘random loneliness tree’ is in fact a more appropriate name for our model.

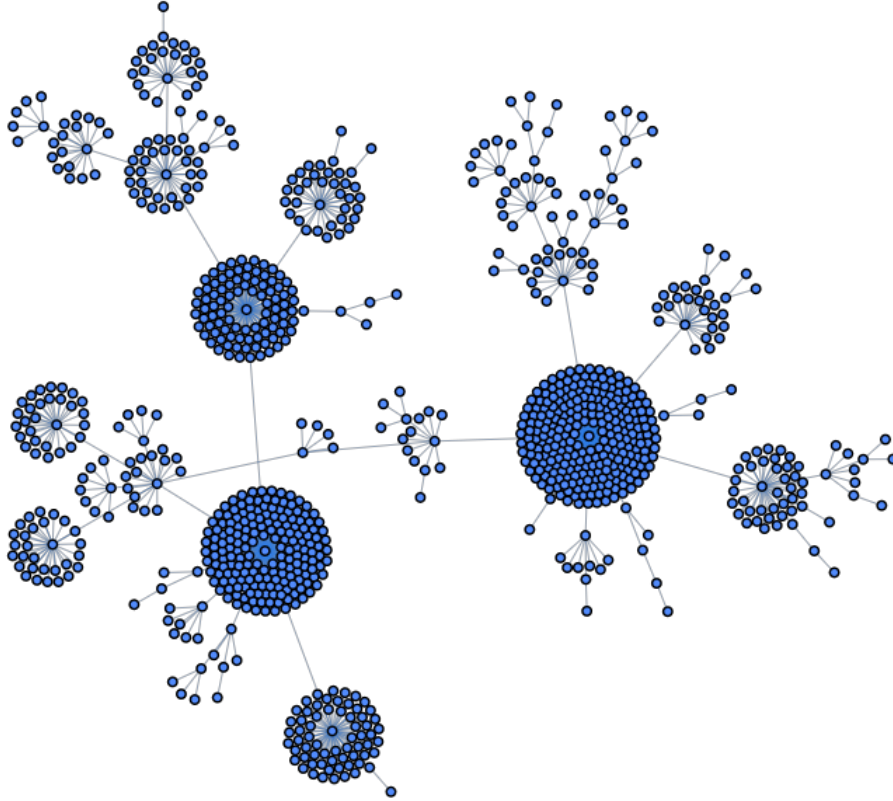


Figure 2: A realisation of  $T_n$  with  $n = 1000$ .

neighbours of  $v$  in  $T_n$ . Define the random variable

$$Z_v := \liminf_{n \rightarrow \infty} \frac{D_n(v)}{n}.$$

A vertex  $v \in \mathbb{N}$  is called a *hub* if  $Z_v > 0$ , that is, if the degree of  $v$  is of linear order asymptotically. A vertex  $w$  is said to be a *child* of vertex  $v$  in  $T_n$  if  $w \in \mathcal{N}(v; T_n)$  and  $v < w$ . If  $w$  is a child of  $v$ , then  $v$  is the *parent* of  $w$ . For  $i, j \in [n]$ , denote by  $d_n(i, j)$  the graph distance between vertices  $i$  and  $j$  in  $T_n$ . We also introduce  $\text{Diam}_n := \max_{i, j \in [n]} d_n(i, j)$ , the diameter of  $T_n$ , and let  $M_n := \max_{i \leq n} \min_{\ell: D_n(\ell)=1} d_n(i, \ell)$  be the maximal distance of any vertex of  $T_n$  to its nearest leaf. For integers  $n, k \geq 1$  we let  $X_n^k = \#\{v \in [n] : D_n(v) = k\}$  be the number of vertices of degree  $k$  in  $T_n$ , and let  $X_n^{\geq k} = \sum_{j \geq k} X_n^j$ .

For any sequence  $(x_n)_{n \geq 1}$ , we define, for all  $n \geq 1$ ,  $\Delta x_n := x_{n+1} - x_n$ . For non-negative  $(x_n)_{n \geq 1}$  and positive  $(y_n)_{n \geq 1}$  we write  $x_n = O(y_n)$  and  $y_n = \Omega(x_n)$  if  $\limsup_{n \rightarrow \infty} \frac{x_n}{y_n} < \infty$  and we write  $x_n = o(y_n)$  and  $y_n = \omega(x_n)$  if  $\lim_{n \rightarrow \infty} \frac{x_n}{y_n} = 0$ . We say that  $x_n = \Theta(y_n)$  if  $x_n = O(y_n)$  and  $x_n = \Omega(y_n)$  both hold. We also use this notation with a  $p$  subscript meaning that the property holds in probability. So, for sequences  $(X_n)$  and  $(Y_n)$  of non-negative random variables, we write  $X_n = O_p(Y_n)$  and  $Y_n = \Omega_p(X_n)$  if for all  $\varepsilon > 0$ , there exist  $M > 0$  and  $K > 0$  such that for all  $n \geq M$ ,  $\mathbf{P}\{X_n > KY_n\} < \varepsilon$ . We write  $X_n = o_p(Y_n)$  and  $Y_n = \omega_p(X_n)$  if for all  $\varepsilon > 0, \delta > 0$ , there exists  $M > 0$  such that for all  $n \geq M$ ,  $\mathbf{P}\{X_n > \delta Y_n\} < \varepsilon$ . We write  $X_n = \Theta_p(Y_n)$  if both  $X_n = O_p(Y_n)$  and  $X_n = \Omega_p(Y_n)$ .

### 3 Local results

In this section we state our results regarding the properties of individual vertices and their close neighbourhoods. First, we state a convergence result for the normalised degree.

**Theorem 3.1** (Convergence of normalised degree). For vertex  $u \in \mathbb{N}$ , the random variables  $D_n(u)$ ,  $L_n(u)$  and  $Z_u$ , defined in Section 2, are such that

$$\frac{D_n(u)}{n} \rightarrow Z_u \text{ and } \frac{L_n(u)}{n} \rightarrow Z_u$$

almost surely as  $n \rightarrow \infty$ .

We observe that  $\sum_{i \geq 1} Z_i \leq 1$  almost surely. Indeed, on the event that this sum exceeds 1, then there must be  $k \in \mathbb{N}$  and  $\delta > 0$  such that  $Z_1 + \dots + Z_k = 1 + \delta$ . But this would imply that there is a finite  $n$  such that the number of leaves neighbouring vertices  $1, \dots, k$  in  $T_n$  satisfies  $L_n(1) + \dots + L_n(k) \geq (1 + \delta/2)n$ . However, the number of leaves in  $T_n$  is at most  $n - 1$  deterministically, so this gives a contradiction.

We conjecture that  $\sum_{i \geq 1} Z_i = 1$  almost surely, implying that, asymptotically, all but a negligible proportion of the vertices are a leaf next to a hub. We discuss this and other open questions on the law of  $(Z_i)_{i \geq 1}$  in Section 7.

A striking property of the friend tree concerns the degree of an edge.

**Theorem 3.2** (Abundance of hubs). The degree of every edge is almost surely asymptotically linear. That is, for any  $m \geq 1$ , if  $W_m$  is the parent of  $m + 1$ , then  $Z_{m+1} + Z_{W_m} > 0$  almost surely.

We also study the limit of the proportion of vertices that are adjacent to  $W_n$ . This theorem shows that the mean of the empirical law of  $D_n(W_n)/n$  given  $T_n$  converges almost surely, implying that this global property of  $T_n$  ‘stabilizes’ as  $n$  grows large. We conjecture that, in fact, the empirical law itself converges almost surely to  $\sum_{i \geq 0} Z_i \delta_{Z_i}$  with respect to the Prokhorov topology. This is in fact equivalent to the conjecture that  $\sum_{i \geq 1} Z_i = 1$  almost surely.

**Theorem 3.3** (Expected degree of  $W_n$ ). As  $n \rightarrow \infty$ ,  $\frac{1}{n} \mathbf{E}[D_n(W_n) | T_n]$  has a positive almost sure limit. Moreover,  $\frac{1}{n} \mathbf{E}[D_n(W_n)]$  converges to some positive number as  $n \rightarrow \infty$ .

Another notable property of random friend trees concerns the probability of a vertex having bounded degree.

**Theorem 3.4** (Bounded degree). Let  $k$  be a positive integer. Fix  $v \in \mathbb{N}$ . Then, for all  $n \geq v$ ,

$$\mathbf{P}\{D_{n+j}(v) = k \ \forall j \geq 0 \mid T_n\} > c_k \mathbf{1}_{[D_n(v)=k]},$$

where the constant  $c_k > 0$  only depends on  $k$ .

In particular, Theorem 3.4 implies that it is impossible to ‘diagnose’ which vertices are the hubs, even at a very large time. Indeed, every vertex, no matter how large its degree is, has probability bounded away from zero to never acquire any new neighbours. We prove Theorems 3.1 and 3.2 in Section 5.1, the proof of Theorem 3.3 can be found in Section 5.2 and Theorem 3.4 is proven in Section 5.3.

## 4 Global results

We now present our results on the global properties of random friend trees.

**Theorem 4.1** (Typical distances). For  $U_n$  and  $V_n$  two uniformly random vertices in  $T_n$ , it holds that the distance between  $U_n$  and  $V_n$  is equal to 2 with probability bounded away from zero. Moreover, the distance between vertex 1 and  $U_n$  is at most 2 with probability bounded away from 0.

The previous result sets the random friend tree apart from the “universality class” of logarithmic trees (and in particular from the uniform random recursive tree and preferential attachment trees), in which typical distances are logarithmic.

The next result shows that, while distances between typical vertices can be very small, the diameter of random friend trees is indeed logarithmic.

**Theorem 4.2** (Diameter of random friend trees). Almost surely

$$1 \leq \liminf_n \frac{\text{Diam}_n}{\log(n)} \leq \limsup_n \frac{\text{Diam}_n}{\log(n)} \leq 4e .$$

In particular, this means that, asymptotically, a path of arbitrary length is present in the tree. Since for every edge at least one of its endpoints is a hub, this implies that an asymptotically unbounded number of hubs are present. We strengthen this statement in Theorem 4.5, where we prove an almost sure polynomial lower bound for the number of hubs in  $T_n$ .

We also study the leaf-depth in  $T_n$ . The next result implies that, although each vertex is at distance at most 1 from a hub, and will therefore eventually be at distance at most two from the nearest leaf, at fixed times there are still exceptional locations in the graph where the nearest leaf is much further away.

**Theorem 4.3** (Leaf depth). Let  $M_n$  be the maximal distance of any vertex to its nearest leaf in  $T_n$ . Then

$$M_n = \Theta \left( \frac{\log(n)}{\log \log(n)} \right) \text{ in probability.}$$

An input to the proof of Theorem 4.3 is the corresponding result for the URRT, which we state as a separate theorem.

**Theorem 4.4** (Leaf depth in URRT). Let  $M'_n$  be the maximal distance of any vertex to its nearest leaf in a URRT. Then,

$$M'_n = \Theta \left( \frac{\log(n)}{\log \log(n)} \right) \text{ in probability.}$$

We prove Theorem 4.3 using Theorem 4.4 and a coupling between URRTs and random friend trees under which distances are at most a factor of two larger in the URRT than in the random friend tree to which it is coupled. This coupling is presented in Lemma 6.3, below.

Another global property of interest is the degree distribution of the tree. We study degree statistics at both ends of the spectrum, for both sub-linear-degree vertices and bounded-degree vertices. We get the following lower bound on the number of hubs in  $T_n$ .

**Theorem 4.5** (Number of hubs). There exists a constant  $\delta > 0.1$  such that

$$\frac{\#\{u \in [n] : Z_u > 0\}}{n^\delta} \rightarrow \infty \text{ a.s.}$$

The following theorem concerns the number of high degree vertices.

**Theorem 4.6** (Abundance of high degree vertices). For any sequence  $(m_n)_{n \geq 1}$ , satisfying  $m_n = o(n)$ , almost surely

$$\lim_{n \rightarrow \infty} X_n^{\geq m_n} = \infty.$$

The next theorem gives polynomial upper and lower bounds for the number of bounded degree vertices.

**Theorem 4.7** (Polynomial bounds on low-degree vertices). There exist constants  $0.1 < \delta \leq \lambda < 0.9$  such that, for any  $k \geq 2$ , almost surely

$$\lim_{n \rightarrow \infty} \frac{X_n^{\geq k}}{n^\delta} = \infty,$$

$$\lim_{n \rightarrow \infty} \frac{X_n^{\geq k}}{n^\lambda} = 0.$$

It has been conjectured by Krapivsky and Redner [15] that a stronger statement is true for the RFT. They conjecture that there exists a constant  $\mu \approx 0.566$  such that,  $n^{-\mu} X_n^{\geq k} \rightarrow X_k$ , where  $X_k$  is a non-degenerate random variable. Moreover, they conjecture that  $X_n^k / X_n^{\geq 2}$  has an almost sure limit which is  $\Theta(k^{-(1+\mu)})$ .

Finally, we show that  $X_n^{\geq k} = \Theta(X_n^{\geq 2})$  almost surely, for any fixed  $k$ .

**Theorem 4.8** (Comparing low-degree vertices). There exists a sequence  $(c_k)_{k \geq 2}$  of positive real numbers, such that for any  $k \geq 2$ ,

$$c_k < \liminf_{n \geq \infty} \frac{X_n^{\geq k+1}}{X_n^{\geq k}} \leq 1$$

almost surely.

We prove Theorem 4.1 in Section 6.1 and we prove Theorem 4.2 in Section 6.2. The proofs of Theorems 4.3 and 4.4 are in Section 6.3. Finally, Theorems 4.5 and 4.6 are proven in Section 6.4 and Theorems 4.7 and 4.8 are proven in Section 6.5. Although Theorem 4.7 is invoked in the proofs of several of the earlier results, we postpone its proof to later in the paper, as it is quite technical.

## 5 Proofs of local properties

### 5.1 Hubs

[Proof of Theorems 3.1 and 3.2] We prove Theorem 3.1 with a submartingale argument, deferring a crucial step of the proof to Section 6.

*Proof of Theorem 3.1.* Fix  $v \in T_n$ . Note that  $L_{n+1}(v) = L_n(v) + 1$  if  $W_n = v$ . For  $V_n = u$  a neighbour of  $v$ ,  $\mathbf{P}\{W_n = v | V_n = u\} = 1/D_n(u)$ . Since  $V_n$  is a uniform sample from  $[n]$ ,



it follows that

$$\mathbf{P}\{L_{n+1}(n) = L_n(v) + 1 \mid T_n\} = \sum_{u \in \mathcal{N}(v; T_n)} \frac{1}{n} \frac{1}{D_n(u)}.$$

Next,  $L_{n+1}(v) = L_n(v) - 1$  if  $V_n = v$  and  $W_n \in \mathcal{L}(v; T_n)$ . Thus,

$$\mathbf{E}[L_{n+1}(v) \mid T_n] = L_n(v) - \frac{1}{n} \frac{L_n(v)}{D_n(v)} + \sum_{u \in \mathcal{N}(v; T_n)} \frac{1}{n} \frac{1}{D_n(u)}.$$

Using that  $L_n(v) \leq D_n(v)$  and, for a leaf  $v$ ,  $D_n(v) = 1$ , we can lower bound the above by  $L_n(v) - \frac{1}{n} + \frac{L_n(v)}{n}$ . By rearrangement it follows that

$$\mathbf{E}\left[\frac{L_{n+1}(v) - 1}{n+1} \mid T_n\right] \geq \frac{L_n(v) - 1}{n}.$$

Thus, for any  $m \in \mathbb{N}$ , the process  $((L_n(v) - 1)/n, n \geq m - 1)$  is a submartingale relative to the filtration generated by the random friend tree process. It is bounded, so it converges almost surely. Furthermore, by the trivial inequalities

$$\frac{L_n(u)}{n} \leq \frac{D_n(u)}{n} \leq \frac{L_n(u) + X_n^{\geq 2}}{n},$$

the joint convergence follows from Theorem 4.7 below, stating that  $n^{-1}X_n^{\geq 2} \rightarrow 0$  almost surely.  $\square$

Recall that a vertex  $u$  is a hub if  $Z_u > 0$ . Theorem 3.2 implies that each edge has at least one endpoint that is a hub, which, in particular, shows that hubs exist.

*Proof of Theorem 3.2.* Fix  $m \in \mathbb{N}$ . For  $n \geq m + 1$ , write  $D_n := D_n(m + 1) + D_n(W_m)$  for the total number of neighbours of vertex  $m + 1$  and vertex  $W_m$  at time  $n$ . Write  $L_n := L_n(m + 1) + L_n(W_m)$  for the number of those neighbours that are leaves. We have  $L_{m+1} \geq 1$  and  $D_{m+1} \geq 3$ . Note that  $D_n$  is non-decreasing and that if  $D_{n+1} = D_n + 1$  then also  $L_{n+1} = L_n + 1$ , so  $\Delta(L_n, D_n) \in \{(1, 1), (0, 0), (-1, 0)\}$ .

Moreover,

$$\mathbf{P}\{\Delta(L_n, D_n) = (1, 1) \mid (L_i, D_i), m + 1 \leq i \leq n\} \geq \frac{1}{n} \left(L_n + \frac{1}{D_n}\right), \quad (5.1)$$

since, to have  $\Delta(L_n, D_n) = (1, 1)$ , it suffices that either  $V_n \in \mathcal{L}(m + 1; T_n) \cup \mathcal{L}(W_m; T_n)$  or else that  $\{V_n, W_n\} = \{m + 1, W_m\}$ . We also have

$$\mathbf{P}\{\Delta(L_n, D_n) = (-1, 0) \mid (L_i, D_i), m + 1 \leq i \leq n\} \leq \frac{\min(2, L_n)}{n}, \quad (5.2)$$

since if  $\Delta(L_n, D_n) = (-1, 0)$ , then  $V_n \in \{m, W_m\}$  and  $W_n \in \mathcal{L}(m; T_n) \cup \mathcal{L}(W_m; T_n)$ .

We claim that  $D_n \rightarrow \infty$  almost surely. We fix  $k \in \mathbb{N}$  and show that the stopping time  $\tau_k = \min\{n \geq m + 1 : D_n \geq k\}$  is finite almost surely. On the event that  $\tau_k < n$ , (5.1) implies that  $\Delta D_n \mid T_n$  stochastically dominates a Bernoulli( $\frac{1}{nk}$ ) random variable. If  $\mathcal{B}_i$  are independent Bernoulli( $\frac{1}{ik}$ ) random variables,  $\sum_{i=m+1}^n \mathcal{B}_i \rightarrow \infty$  almost surely, so  $\tau_k < \infty$  almost surely. It then follows that  $\#\{n : L_n > 0\} = \infty$  almost surely. Indeed,  $D_n \rightarrow \infty$  implies that  $\Delta(D_n, L_n) = (1, 1)$  infinitely many times, so if  $L_n$  ever hits zero it almost surely becomes positive again.

Let  $(J_k, k \geq 0)$  be the sequence of jump times of the process  $(\Delta(L_n, D_n))_n$ , that is  $J_0 = m + 1$  and  $J_k = \min\{\ell > J_{k-1} : \Delta(L_{\ell-1}, D_{\ell-1}) \neq (0, 0)\}$  for  $k \geq 1$ . This sequence has infinite length because  $D_n \rightarrow \infty$  almost surely.

Our goal is to show that  $L_n$  grows linearly. To do so, we couple  $L_n$  to an urn process. Inequalities (5.1) and (5.2) suggest that the growth of  $L_n$  is similar to the growth of the number of black balls in a standard Pólya urn with black and white balls. One difference being that, at time  $n$ , with probability at most  $2/n$ , a black ball is replaced by a white ball. Nonetheless, we exhibit a coupling between  $L_n$  and the number of black balls in a standard Pólya urn of black and white balls of size  $n$ . More precisely, a coupling where  $L_n$  is greater than the number of black balls in the Pólya urn of size  $n$ . This coupling can fail, meaning that it is valid until a random time  $S$  that is finite with positive probability. We say that the coupling succeeds if  $S = \infty$ . We show that this coupling succeeds with probability greater than 0, and that, if it fails, we may try again by starting a new coupling at a subsequent time. This guarantees that one of the coupling attempts is successful, proving that  $L_n$  grows linearly, because the number of black balls in a standard Pólya urn grows linearly almost surely.

The coupling is started at a time where  $L_n$  is at least 10. The bounds (5.1) and (5.2) on the transition probabilities show that the process  $(L_{J_k}, k \geq 0)$  stochastically dominates a simple symmetric random walk reflected at 0. This, in particular, implies that there are infinitely many  $n$  such that  $L_n \geq 10$ , because  $J_k \rightarrow \infty$  almost surely.

Let  $\rho_1$  be the first time for which  $L_{\rho_1} \geq 10$  (and so  $L_{\rho_1} = 10$ ). A first coupling is started from time  $\rho_1$ . For any  $k$ , if the  $k$ th coupling fails, let  $\rho_{k+1}$  be the first time after the failure at which  $L_{\rho_{k+1}} \geq 10$  and start the  $(k+1)$ st coupling from that time. We show that there is a  $c > 0$  so that for each  $k$ , given that couplings  $1, \dots, k-1$  all failed, the  $k$ th coupling succeeds with probability at least  $c$ . This implies that there is an almost surely finite  $M$  so that the  $M$ th coupling succeeds.

So let us fix some  $N > 0$  and condition on  $\rho_k = N$ . Now define  $B_n$  the number of black balls in a Pólya urn starting at time  $N$  with 5 black balls and  $N-5$  white balls (with the standard replacement rule that a drawn ball is replaced along with one extra ball of the same colour). Note that  $L_N = L_{\rho_k} \geq 10$  so  $L_N - B_N \geq 5$ . We couple  $L_n$  and  $B_n$  from time  $N$  onwards and we say the coupling fails at time  $S$  if  $S$  is the first time  $S > N$  at which  $L_S - B_S \leq 4$ . If the coupling never fails we set  $S = \infty$ . Then, for  $n \in [N, S]$ , we can couple  $L_n$  and  $B_n$  such that if  $B_{n+1} = B_n + 1$  then  $L_{n+1} = L_n + 1$ . From this coupling, for  $n \geq N$ ,

$$\begin{aligned} \mathbf{P} \{ \Delta(L_n - B_n) = 1 \mid S > n \} &\geq \frac{L_n - B_n}{n} \\ \mathbf{P} \{ \Delta(L_n - B_n) = -1 \mid S > n \} &\leq \frac{2}{n}. \end{aligned}$$

This means that until the coupling fails,  $L_n - B_n$  can be coupled to a symmetric random walk for which an increment with value 1 is twice as likely as an increment with value  $-1$ . We introduce  $R_k$ , a random walk with  $R_0 = 5$  and

$$R_{k+1} - R_k = \begin{cases} +1 & \text{with probability } \frac{2}{3}, \\ -1 & \text{with probability } \frac{1}{3}. \end{cases}$$

Set  $I_0 = N$  and let  $I_{k+1} = \min\{j \geq I_k : \Delta(L_n - B_n) \neq 0\}$  be the  $k$ th jump time of  $L_n - B_n$ . Then  $(R_k, k \geq 0)$  and  $(L_n - B_n, n \geq N)$  can be coupled such that if  $I_k \leq S$

$$L_{I_k} - B_{I_k} \geq R_k.$$

With positive probability  $R_k > 4$  for all  $k$ , so with positive probability, not depending on  $N$ ,  $S = \infty$ . This shows that, almost surely, one of the coupling attempt succeeds. Suppose that the  $k$ th coupling succeeds and that  $\rho_k = N$ . Then, for  $B_n$  as above,  $L_n \geq B_n$  for  $n \geq N$ , so since

$$\lim_{n \rightarrow \infty} \frac{B_n}{n} > 0 \text{ almost surely,}$$

by a standard result on Pólya urns (see Mahmoud [17, Section 3.2]), we also get that

$$\liminf_{n \rightarrow \infty} \frac{L_n}{n} > 0 \text{ almost surely,}$$

which implies the statement.  $\square$

## 5.2 Expected degree of $W_n$

[Proof of Theorem 3.3]

We show the statement using the almost sure martingale convergence theorem by identifying a supermartingale. Set  $Y_n := \mathbf{E}[D_n(W_n) | T_n]$  so that  $Y_n$  is adapted to  $\sigma(T_n)$  and note that

$$Y_n = \frac{1}{n} \sum_{i \in [n]} \mathbf{E}[D_n(W_n) | T_n, V_n = i] = \frac{1}{n} \sum_{i \in [n]} \left( \frac{1}{D_n(i)} \sum_{j \sim_{T_n} i} D_n(j) \right).$$

To identify the supermartingale, we study

$$\mathbf{E}[(n+1)Y_{n+1} | T_n] = \mathbf{E} \left[ \sum_{i \in [n+1]} \sum_{j \sim_{T_{n+1}} i} \frac{D_{n+1}(j)}{D_{n+1}(i)} \middle| T_n \right].$$

Observe that the only randomness in  $Y_{n+1}$ , conditional on  $T_n$ , comes from the choice of  $W_n$ . It holds that  $D_{n+1}(W_n) = D_n(W_n) + 1$ , and the new neighbour of  $W_n$  (vertex  $n+1$ ) has degree 1. Moreover,  $D_{n+1}(n+1) = 1$ , because every vertex starts as a leaf, and its neighbour is  $W_n$ . Finally,  $D_{n+1}(i) = D_n(i)$  for all other  $i$ . Therefore, we get the following equalities for the different terms in  $\sum_{i \in [n+1]} \sum_{j \sim_{T_{n+1}} i} \frac{D_{n+1}(j)}{D_{n+1}(i)}$ :

$$\sum_{j \sim_{T_{n+1}} i} \frac{D_{n+1}(j)}{D_{n+1}(i)} = \begin{cases} \frac{1}{D_n(W_n)+1} + \sum_{j \sim_{T_n} W_n} \frac{D_n(j)}{D_n(W_n)+1} & \text{for } i = W_n \\ D_n(W_n) + 1 & \text{for } i = n+1 \\ \frac{\mathbf{1}_{\{i \sim_{T_n} W_n\}}}{D_n(i)} + \sum_{j \sim_{T_n} i} \frac{D_n(j)}{D_n(i)} & \text{otherwise.} \end{cases}$$

Combining these cases, we see that

$$\begin{aligned} & \mathbf{E}[(n+1)Y_{n+1} | T_n] \\ &= \sum_{i \in [n]} \sum_{j \sim_{T_n} i} \frac{D_n(j)}{D_n(i)} \\ &+ \mathbf{E} \left[ \frac{1}{D_n(W_n)+1} - \left( \frac{1}{D_n(W_n)} - \frac{1}{D_n(W_n)+1} \right) \sum_{j \sim_{T_n} W_n} D_n(j) \middle| T_n \right] \\ &+ \mathbf{E} \left[ D_n(W_n) + 1 + \sum_{j \sim_{T_n} W_n} \frac{1}{D_n(j)} \middle| T_n \right]. \end{aligned}$$

Then, using that  $D_n(j) \geq 1$  for all  $j \sim_{T_n} i$  we get that the second term on the right hand side is positive. To get an upper bound for the third term, we again use that  $D_n(j) \geq 1$  to get that

$$\mathbf{E}[(n+1)Y_{n+1} | T_n] \leq \sum_{i \in [n]} \sum_{j \sim_{T_n} i} \frac{D_n(j)}{D_n(i)} + \mathbf{E}[2D_n(W_n) + 1 | T_n] \leq (n+2)Y_n + 1,$$

so that  $\mathbf{E}[Y_{n+1}/(n+2) \mid Y_n] \leq Y_n/(n+1) + 1/(n+1)^2$ , and therefore  $Y_n/(n+1) - \sum_{i=1}^n 1/i^2$  is a supermartingale in the filtration generated by  $T_n$ , and therefore has an almost sure limit. Since  $1/i^2$  is summable, it follows that  $\frac{1}{n}\mathbf{E}[D_n(W_n) \mid T_n]$  has an almost sure limit.

To see that the limit is positive, note that Theorem 3.2 implies that for any  $\varepsilon > 0$ , there is a  $\delta > 0$  so that  $\mathbf{P}\{Z_1 + Z_2 > \delta\} > 1 - \varepsilon$ .

Then, observe that

$$\begin{aligned} \mathbf{E}\left[\frac{1}{n}D_n(W_n) \mid T_n\right] &> \frac{1}{n}D_n(1)\mathbf{P}\{W_n = 1 \mid T_n\} + \frac{1}{n}D_n(2)\mathbf{P}\{W_n = 2 \mid T_n\} \\ &\geq \frac{L_n(1)^2 + L_n(2)^2}{n^2}, \end{aligned}$$

since  $D_n(i) \geq L_n(i)$  and  $\mathbf{P}\{W_n = i \mid T_n\} \geq L_n(i)/n$  for  $i = 1, 2$ . Then, note that, if  $L_n(1) + L_n(2) > \delta n/2$  then

$$\frac{L_n(1)^2 + L_n(2)^2}{n^2} \geq (\max\{L_n(1)/n, L_n(2)/n\})^2 \geq (\delta/4)^2,$$

so

$$\mathbf{P}\left\{\frac{1}{n}\mathbf{E}[D_n(W_n) \mid T_n] > \delta^2/16\right\} \geq \mathbf{P}\{L_n(1) + L_n(2) > \delta n/2\}.$$

But,  $\frac{1}{n}(L_n(1) + L_n(2)) \rightarrow Z_1 + Z_2$  almost surely, so

$$\liminf_{n \rightarrow \infty} \mathbf{P}\{L_n(1) + L_n(2) > \delta n/2\} > 1 - \varepsilon$$

so also

$$\liminf_{n \rightarrow \infty} \mathbf{P}\left\{\frac{1}{n}\mathbf{E}[D_n(W_n) \mid T_n] > \delta^2/16\right\} > 1 - \varepsilon,$$

which implies the statement.

The convergence in expectation follows from the bounded convergence theorem, since  $\frac{1}{n}D_n(W_n) \leq 1$  deterministically.

### 5.3 Eternal leaves and eternal degree $k$ vertices

[Proof of Theorem 3.4] Note that, for any integer  $i$ ,  $D_n(i)$  is increasing in  $n$  and therefore it has an almost sure limit (that might be infinite). For a vertex  $\ell$  that is a leaf at time  $m$ , we say it is *temporary* if  $\lim_{n \rightarrow \infty} D_n(\ell) > D_m(\ell) = 1$ . Otherwise we call it *eternal*. Similarly, we call a vertex  $v$  that has degree  $k$  at time  $m$  *temporary* if  $\lim_{n \rightarrow \infty} D_n(v) > D_m(v) = k$  and otherwise we call it *eternal*. Informally, our next proposition says that, only a bounded number of leaves next to a given hub ever stop being a leaf.

**Proposition 5.1.** For  $n \geq v$ , let  $S_n(v)$  be the number of temporary leaves attached to  $v$  at time  $n$ . If  $\mathbf{P}\{Z_v > 0\} > 0$ , then conditional on  $Z_v > 0$ ,  $S_n(v)$  is tight, that is, for all  $\varepsilon > 0$  there exists a constant  $M > 0$  such that  $\mathbf{P}\{S_n(v) > M \mid Z_v > 0\} < \varepsilon$  for all  $n \geq v$ .

*Proof.* Fix  $\varepsilon > 0$ . Suppose  $v$  is a hub, that is  $Z_v = \lim_{n \rightarrow \infty} \frac{D_n(v)}{n} > 0$ . This implies that there is a  $\delta$  and a  $N > v$  such that  $\mathbf{P}\{\forall n \geq N L_n(v) > \delta n \mid Z_v > 0\} > 1 - \varepsilon/2$ . We show that there exists a constant  $K$  such that  $\mathbf{E}[S_k(v)\mathbf{1}_{\{\forall n \geq N L_n(v) > \delta n\}}] < K$  for all  $k \geq N$ . We first show that this implies the statement. Write  $\mathbf{P}\{Z_v > 0\} = \rho$ , so that  $\mathbf{E}[S_k(v)\mathbf{1}_{\{\forall n \geq N L_n(v) > \delta n\}} \mid Z_v > 0] < K/\rho$ . Observe that  $L_n(v) > \delta n \forall n \geq N$  implies  $Z_v > 0$ . Then, we see that for  $k \geq N$  and  $\varepsilon > 0$ ,

$$\begin{aligned}
 & \mathbf{P} \left\{ S_k(v) > \frac{2K}{\rho\varepsilon} \mid Z_v > 0 \right\} \\
 & \leq \mathbf{P} \left\{ \exists n \geq N : L_n(v) \leq \delta n \mid Z_v > 0 \right\} + \mathbf{P} \left\{ S_k(v) > \frac{2K}{\rho\varepsilon}, \forall n \geq N L_n(v) > \delta n \mid Z_v > 0 \right\} \\
 & \leq \varepsilon/2 + \mathbf{P} \left\{ S_k(v) > \frac{2K}{\rho\varepsilon} \mid \forall n \geq N L_n(v) > \delta n \right\} \mathbf{P} \left\{ L_n(v) > \delta n \forall n \geq N \mid Z_v > 0 \right\} \\
 & \leq \varepsilon/2 + \frac{\varepsilon\rho}{2K} \mathbf{E} \left[ S_k(v) \mid \forall n \geq N L_n(v) > \delta n \right] \mathbf{P} \left\{ L_n(v) > \delta n \forall n \geq N \mid Z_v > 0 \right\} \\
 & = \varepsilon/2 + \frac{\varepsilon\rho}{2K} \mathbf{E} \left[ S_k(v) \mathbf{1}_{[\forall n \geq N L_n(v) > \delta n]} \mid Z_v > 0 \right] < \varepsilon,
 \end{aligned}$$

where we use Markov's inequality in the penultimate line.

We now show that there exists a constant  $K$  such that  $\mathbf{E} [S_k(v) \mathbf{1}_{[\forall n \geq N L_n(v) > \delta n]}] < K$  for all  $k \geq N$ . Note that, at any time  $M \geq k$ , if  $w$  is a leaf neighbouring  $v$ , then  $w$  stops being a leaf if  $V_M = v$  and  $W_M = w$ . Conditionally on  $T_M$ , this occurs with probability  $1/(MD_M(v))$ , so the probability that this happens for some vertex in  $\mathcal{L}(v; T_k)$  is at most  $k/(MD_M(v))$ , because  $|\mathcal{L}(v; T_k)| < k$ . Therefore, the probability that the number of leaves in  $\mathcal{L}(v; T_k)$  that are no longer leaves increases at time  $M \geq k$  satisfies

$$\begin{aligned}
 & \mathbf{P} \left\{ \Delta |\mathcal{L}(v; T_k) \cap \mathcal{L}(v; T_M)^c| = 1, L_n(v) > \delta n \forall n \geq N \right\} \\
 & \leq \mathbf{P} \left\{ \Delta |\mathcal{L}(v; T_k) \cap \mathcal{L}(v; T_M)^c| = 1, L_M(v) > \delta M \right\} \leq \frac{k}{\delta M^2}.
 \end{aligned}$$

Therefore

$$\mathbf{E} [S_k(v) \mathbf{1}_{[L_n(v) > \delta n \forall n \geq N]}] \leq \sum_{M \geq k} \frac{k}{\delta M^2} \leq K$$

for some constant  $K$  not depending on  $k$ , which proves the claim.  $\square$

The next corollary, stating that the number of eternal leaves attached to any edge grows linearly with high probability, is a consequence of the proposition above and Theorem 3.2. Indeed, since every edge has linear degree almost surely, given the presence of edge  $(u, v)$ , either  $u$  or  $v$  has probability at least  $1/2$  of being a hub.

**Corollary 5.2.** Fix an edge  $(u, v)$  and for  $n \geq \max(u, v)$  define  $E_n(u, v) := |\{w \in \mathcal{N}(u; T_n) \cup \mathcal{N}(v; T_n) : w \text{ is an eternal leaf}\}|$ . Then,  $E_n(u, v)$  grows linearly with high probability, that is, there exists  $\delta > 0$  such that for every  $\varepsilon > 0$  there is a  $N = \mathcal{N}(u, v)$  such that

$$\mathbf{P} \left\{ \exists n > N : E_n(u, v) < \delta n \right\} < \varepsilon.$$

We now prove Theorem 3.4.

*Proof of Theorem 3.4.* Fix  $v \in \mathbb{N}$ . We first prove the statement for  $k = 1$ , and then discuss how to adapt the proof to general  $k$ . If  $D_n(v) = 1$  then let  $w$  denote the unique neighbour of  $v$  in  $T_n$ . We show that  $v$  is an eternal leaf with positive probability by showing that, with positive probability, the vertex  $w$  acquires a large number of leaf neighbours, ensuring that both the degree of  $w$  grows and that, when a new vertex is attached to a uniform neighbour of  $w$ , it is unlikely that  $v$  is chosen. Fix  $N$  and let

$$\tau = \min \{t \in \mathbb{N} : \#\{n < i \leq t : V_i \in \{v, w\}\} = N\}$$

be the random time at which a new vertex is attached to a random neighbour of either  $v$  or  $w$  exactly  $N$  times since time  $n$ . Define  $A_N$  as the event  $\{\#\{n < i \leq \tau : V_i = v\} = N\}$ . Since  $V_i$  is chosen uniformly at random, with probability  $2^{-N}$ ,  $V_i = v$  exactly  $N$  times between times  $n$  and  $\tau$ . So,  $\mathbf{P} \{A_N\} = 2^{-N}$ . Recall that  $D_n(v) = 1$ , so conditionally on  $A_N$ ,  $D_\tau(v) = 1$  because for all  $i \in [n, \tau]$ ,  $V_i \neq u$ , and thus  $W_i \neq v$ . Moreover,  $L_\tau(w) \geq N$ ,

because, conditioned on  $A_N$ , when  $V_i = v$ ,  $W_i = w$ , so that a new leaf is attached to  $w$ . Since for all  $i \in [n, \tau]$   $V_i \neq u$ , these leaves stay leaves until time  $\tau$ . We show that, on the event  $A_N$ ,  $v$  is an eternal leaf with positive probability, because it is likely that  $w$  continues to acquire many leaf neighbours beyond time  $\tau$ , making it unlikely for the degree of  $v$  to grow.

For  $j \geq 0$ , set  $a_j = \tau \cdot 2^j$  and  $\ell_j = N(5/4)^j$  so that  $a_{j+1} - a_j = a_j$  and  $\ell_{j+1} - \ell_j = \ell_j/4$ . Define the following events for  $j \geq 0$ :

$$\begin{aligned} E_j &:= \{D_{a_{j+1}}(v) > 1\}, \\ F_j &:= \{\#\{i \in (a_j, a_{j+1}] : V_i = w\} > m(j+1)\}, \\ G_j &:= \{L_{a_{j+1}}(w) < \ell_{j+1}\}, \\ B_j &:= E_j \cup F_j \cup G_j, \end{aligned}$$

where  $m > 0$  is such that  $m < N/20$ . In words,  $E_j$  is the event that vertex  $v$  is not a leaf in  $T_{a_{j+1}}$ . The event  $F_j$  corresponds to the event that between steps  $(a_j, a_{j+1}]$ , new vertices attach to a random neighbour of  $w$  more than  $m(j+1)$  times and  $G_j$  corresponds to the event that  $w$  has less than  $\ell_{j+1}$  leaf neighbours in  $T_{a_{j+1}}$ . It suffices to prove that

$$\mathbf{P} \left\{ \bigcup_{j \geq 0} B_j \mid A_N \right\} < 3/4. \quad (5.3)$$

Indeed,

$$\mathbf{P} \{v \text{ is an eternal leaf}\} = \mathbf{P} \left\{ \bigcap_{j \geq 0} E_j^c \right\} \geq \mathbf{P} \left\{ \bigcap_{j \geq 0} B_j^c \right\},$$

since  $B_j^c \subset E_j^c$ . By (5.3),

$$\mathbf{P} \left\{ \bigcap_{j \geq 0} B_j^c \right\} \geq \mathbf{P} \left\{ \bigcap_{j \geq 0} B_j^c, A_N \right\} = \mathbf{P} \{A_N\} \mathbf{P} \left\{ \bigcap_{j \geq 0} B_j^c \mid A_N \right\} \geq \frac{1}{4} 2^{-N}.$$

This implies

$$\mathbf{P} \{v \text{ is an eternal leaf}\} \geq 2^{-(N+2)},$$

proving Theorem 3.4 for  $k = 1$ . To show (5.3) we begin by noting that

$$\mathbf{P} \left\{ E_j \mid \bigcap_{i=0}^{j-1} B_i^c, A_N \right\} \leq \mathbf{E} \left[ \mathbf{E} \left[ \sum_{i=a_j}^{a_{j+1}} \frac{1}{i \ell_j} \mid \tau \right] \right] \leq \frac{1}{\ell_j} (1 + \log(2)),$$

where the first inequality holds since the conditioning implies that vertex  $w$  has degree at least  $\ell_j$  from time  $a_j$  onwards. Next, since the probability of  $V_i = w$  equals to  $1/i \leq 1/a_j$  for all  $i \geq a_j$ , we deduce that, for  $j \geq 1$ ,

$$\mathbf{P} \left\{ F_j \mid \bigcap_{i=0}^{j-1} B_i^c, A_N \right\} \leq \mathbf{P} \left\{ \text{Bin}(a_{j+1} - a_j, 1/a_j) > m(j+1) \right\} \leq e^{-m(j+1)/3},$$

by a Chernoff bound ([13, Theorem 2.1.]). Lastly, under the previous conditioning and given  $F_j^c$ , the probability of  $G_j$  is less than the probability of creating fewer than  $\ell_{j+1} + m(j+1) - \ell_j$  new leaf neighbours for  $w$  between steps  $(a_j, a_{j+1}]$ . Conditionally on  $F_j^c$ , the probability of attaching vertex  $i+1$  to  $w$  at time  $i \in (a_j, a_{j+1}]$  is at least  $(\ell_j - m(j+1))/a_{j+1}$ . Thus

$$\begin{aligned} \mathbf{P} \left\{ G_j \mid \bigcap_{i=0}^{j-1} B_i^c, F_j^c, A_N \right\} &\leq \mathbf{P} \left\{ \text{Bin} \left( a_{j+1} - a_j, \frac{\ell_j - m(j+1)}{a_{j+1}} \right) \leq \ell_{j+1} + m(j+1) - \ell_j \right\} \\ &\leq e^{-\ell_j/90}, \end{aligned}$$

where we use that  $m < N/20$ , so that  $m(j+1) \leq \frac{1}{10}\ell_j$  and  $\ell_j - m(j+1) > 0$  for all  $j \geq 0$ , and the bound then follows from the Chernoff bound. Putting everything together gives us that,

$$\mathbf{P} \left\{ B_j \mid \bigcap_{i=0}^{j-1} B_i^c, A_N \right\} \leq \ell_j^{-1}(1 + \log(2)) + e^{-m(j+1)/3} + e^{-\ell_j/90},$$

and

$$\begin{aligned} \mathbf{P} \left\{ \bigcup_{j \geq 0} B_j \mid A_N \right\} &\leq \sum_{j \geq 0} \mathbf{P} \left\{ B_j \mid \bigcap_{i=0}^{j-1} B_i^c, A_N \right\} \\ &\leq \frac{1}{N} \sum_{j \geq 0} (4/5)^j (1 + \log(2)) + \sum_{j \geq 1} e^{-m(j+1)/3} + \sum_{j \geq 0} e^{-\mathcal{N}(5/4)^j/90}. \end{aligned}$$

Now, choose  $m$  sufficiently large so that the second sum is smaller than  $1/4$ , then choose  $N$  large enough so that  $m < N/20$  and the first and the third sum are both smaller than  $1/4$ ; this proves the statement for  $k = 1$ . Next, we adapt the statement for general  $k$ . Conditionally on  $T_n$ , if  $D_n(v) = k$  then we define  $\tau$  as the random time at which a new vertex has been attached to a random friend of vertex  $v$  or one of its  $k$  neighbours exactly  $kN$  times. That is,  $\tau = \min\{t : \#\{n < i \leq t : V_i \in \{v\} \cup \mathcal{N}(v; T_n)\} = kN\}$ . Then, with probability at least  $(k+1)^{-kN}$  exactly  $N$  of these  $kN$  new leaves are attached to each of the  $k$  neighbours of  $v$ . We call this event  $A_N = \{\forall w \in \mathcal{N}(v; T_n) : \#\{n < i \leq \tau : V_i = w\} = N\}$ . Conditionally on  $A_N$ , we show that  $v$  is an eternal degree  $k$  vertex with positive probability, because it is likely that all of the neighbours of  $v$  continue to acquire many leaf neighbours beyond time  $\tau$ , making it unlikely for the degree of  $v$  to grow.

Let  $w_1, \dots, w_k$  denote the neighbours of  $v$  in  $T_n$ . Define the events

$$\begin{aligned} E_j^k &:= \{D_{a_{j+1}}(v) > k\}, \\ F_j^k &:= \bigcup_{l=1}^k \{\#\{i \in (a_j, a_{j+1}] : V_i = w_l\} > m(j+1)\}, \\ G_j^k &:= \bigcup_{l=1}^k \{L_{a_{j+1}}(w_l) < \ell_{j+1}\}, \\ B_j^k &:= E_j \cup F_j \cup G_j. \end{aligned}$$

Following the same arguments used in the case of  $k = 1$ , it follows that, choosing  $m$  and  $N$  sufficiently large, we have

$$\mathbf{P} \left\{ \bigcup_{j \geq 0} B_j^k \mid A_N \right\} < 3/4,$$

which implies the statement for general  $k$ . □

Note that it follows from Theorem 3.4 that, for  $E_n^k$  the number of eternal degree  $k$  vertices in  $T_n$ ,

$$\mathbf{E} [E_n^k] > c_k \mathbf{E} [X_n^k]. \tag{5.4}$$

## 6 Proofs of global properties

### 6.1 Typical distances

[Proof of Theorem 4.1]

Theorem 4.1 is a consequence of Theorem 3.2. Indeed, almost surely, one of vertices 1 and 2 is a hub, so a positive proportion of vertices is a neighbour of vertex 1 or of vertex 2. This implies that with probability bounded away from zero, both  $U_n$  and  $V_n$  are a neighbour of either vertex 1 or 2, so that the distance between them is two and the distance from  $U_n$  to 1 is at most two. More formally, by Theorem 3.2 there is an  $i \in \{1, 2\}$  and an  $\epsilon > 0$  so that  $\mathbf{P}\{Z_i > \epsilon\} > \epsilon$ . Then, for  $n$  large enough,  $\mathbf{P}\{D_n(i) > \epsilon n/2\} > \epsilon/2$ . For such  $n$ ,

$$\mathbf{P}\{d_n(U_n, 1) \leq 2\} \geq \mathbf{P}\{D_n(i) > \epsilon n/2, d_n(i, U_n) = 1\} \geq \epsilon^2/4$$

and

$$\mathbf{P}\{d_n(U_n, V_n) = 2\} \geq \mathbf{P}\{D_n(i) > \epsilon n/2, d_n(i, U_n) = 1, d_n(i, V_n) = 1\} \geq \epsilon^3/8,$$

which proves the statement.

## 6.2 Diameter

[Proof of Theorem 4.2] We show that  $\text{Diam}_n$  grows logarithmically almost surely, with explicit asymptotic lower and upper bounds. We start by proving a lower bound.

**Lemma 6.1.** Almost surely

$$\liminf_n \frac{\text{Diam}_n}{\log(n)} \geq 1.$$

*Proof.* Among all the paths of length  $\text{Diam}_n$  present at time  $n$ , let us choose one. Denote it by  $(i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_{\text{Diam}_n})$ . Let us remark that for  $n \geq 3$ ,  $\text{Diam}_n$  is always at least 2. Vertices  $i_1$  and  $i_{\text{Diam}_n-1}$  are such that at most one of their neighbours is not a leaf (otherwise there would be a path of length  $\text{Diam}_n + 1$ ). This implies that, at time  $n$ , conditioned on  $V_n = i_1$ , with probability at least  $1/2$  we have that  $W_n \in \mathcal{L}(i_1; T_n)$  (the same holds for  $i_{\text{Diam}_n-1}$ ). But, if  $W_n \in \mathcal{L}(i_1; T_n) \cup \mathcal{L}(i_{\text{Diam}_n-1}; T_n)$  then the diameter increases by 1. Because  $\mathbf{P}\{V_n \in \{i_1, i_{\text{Diam}_n-1}\} \mid \text{Diam}_n \geq 3\} = 2/n$  and  $\mathbf{P}\{V_n \in \{i_1, i_{\text{Diam}_n-1}\} \mid \text{Diam}_n = 2\} = 1/n$  (note that if  $\text{Diam}_n = 2$ , then  $i_1 = i_{\text{Diam}_n-1}$ ), then

$$\mathbf{E}[\Delta \text{Diam}_n \mid T_n] \geq \frac{1}{n} \mathbf{1}_{[\text{Diam}_n \geq 3]} + \frac{1}{2n} \mathbf{1}_{[\text{Diam}_n = 2]}. \quad (6.1)$$

To prove the lemma, we first need to show that, almost surely,  $\text{Diam}_n$  reaches 3 in finite time. Using (6.1), there exists a coupling between  $\text{Diam}_n$  and

$$S_n := 2 + \sum_{i=4}^n Z_i,$$

where  $Z_i$  are independent Bernoulli random variables with parameter  $1/(2i)$ , such that  $\text{Diam}_n \geq S_n$  for  $n \geq 3$ . Let  $M$  be the first time when  $\text{Diam}_n = 3$  and  $M'$  the first time when  $S_n = 3$ . By our coupling of  $\text{Diam}_n$  and  $S_n$ ,  $M \leq M'$ . A direct application of [4, Exercise 2.9] shows that  $\liminf S_n = \infty$  almost surely and therefore  $M'$  (and in turn  $M$ ) are finite almost surely. We can now introduce a coupling of  $\text{Diam}_n$  from time  $M$  onward. Conditionally on  $M = m$ , (6.1) implies that  $\text{Diam}_n$  can be coupled from  $m$  onward to

$$H_n^{(m)} := 3 + \sum_{i=m}^n X_i,$$

where  $X_i$  are independent Bernoulli random variables with parameters  $1/i$ , such that  $\text{Diam}_n \geq H_n^{(m)}$ . Another direct application of [4, Exercise 2.9] implies that, for fixed  $m$ ,



## Random friend trees

$$\liminf_n \frac{H_n^{(m)}}{\log(n)} \geq 1 \text{ a.s.}$$

By our coupling, if  $M = m$ ,  $\liminf \frac{\text{Diam}_n}{\log(n)} \geq 1$  almost surely. Using that  $M$  is finite almost surely concludes the proof.  $\square$

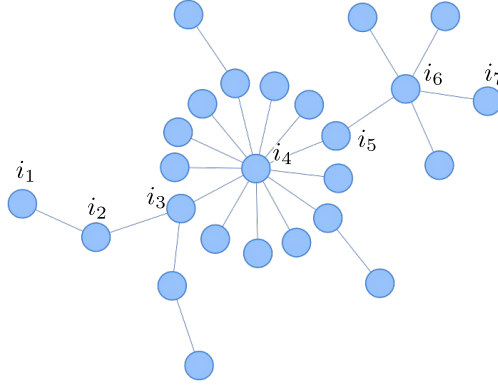


Figure 3: Illustration of a RFT of size 25 and diameter 6, with vertices of one of the paths of length 6 highlighted.

Next, we prove an upper bound for the diameter.

**Lemma 6.2.** Almost surely

$$\limsup_n \frac{\text{Diam}_n}{\log(n)} \leq 4e .$$

To prove this lemma we couple the random friend tree with the URRT process. This coupling is the subject of the following lemma.

**Lemma 6.3.** The random friend tree  $(T_n, n \geq 2)$  and the uniform random recursive tree  $(T'_n, n \geq 2)$  can be coupled in such a way that for any  $i, j \in [n]$ ,

$$d_n(i, j) \leq 2d'_n(i, j) ,$$

where  $d_n(i, j)$  is the graph distance between vertex  $i$  and  $j$  in  $T_n$ , and  $d'_n(i, j)$  is the graph distance between vertex  $i$  and  $j$  in  $T'_n$ .

*Proof.* Note that  $T_2 = T'_2$  so the statement holds for  $n = 2$ . Fix  $m \geq 2$  and suppose that we coupled  $(T_2, \dots, T_m)$  and  $(T'_2, \dots, T'_m)$  such that for all  $i, j \in [m]$ ,  $d_m(i, j) \leq 2d'_m(i, j)$ . Now, sample uniformly at random  $V_m \in [m]$  and let  $W_m$  be a uniform neighbour of  $V_m$  in  $T_m$ . Let  $T_{m+1}$  be the tree obtained by including vertex  $m+1$  and edge  $\{W_m, m+1\}$  in  $T_m$  and let  $T'_{m+1}$  be the tree obtained by including vertex  $m+1$  and edge  $\{V_m, m+1\}$  in  $T'_m$ . Observe that, for  $i, j \in [m]$ ,  $d_{m+1}(i, j) = d_m(i, j)$  and  $d'_{m+1}(i, j) = d'_m(i, j)$ . Now, let  $i \in [m]$  and compute

$$\begin{aligned} d_{m+1}(i, m+1) &\leq d_{m+1}(i, V_m) + d_{m+1}(V_m, m+1) = d_m(i, V_m) + 2 \\ &\leq 2d'_m(i, V_m) + 2 \leq 2d'_m(i, m), \end{aligned}$$

where we use the triangle inequality, the induction hypothesis and the fact that for all  $i \in [m]$ ,  $d'_m(i, m) \leq 1 + d'_m(i, V_m)$ .  $\square$

*Proof of Lemma 6.2.* Couple  $(T_n, n \geq 1)$  to the uniform random recursive tree  $(T'_n, n \geq 1)$  as in Lemma 6.3 and observe that

$$\text{Diam}_n = \max_{i,j \leq n} d_n(i,j) \leq 2 \max_{i,j \leq n} d'_n(i,j) \leq 4 \max_{i \leq n} d'_n(1,i).$$

Moreover, by Corollary 1.3 of Addario-Berry and Ford [1],

$$\frac{\max_{i \leq n} d'_n(1,i)}{\log n} \rightarrow e \text{ almost surely,}$$

which concludes the proof of the lemma.  $\square$

### 6.3 Leaf-depth

[Proof of Theorems 4.3 and 4.4] Theorem 3.2 implies that, asymptotically, almost surely each vertex is at distance at most 1 from a hub. By Theorem 3.1, each hub has mostly leaf neighbours. This suggests that at large times, most vertices that are not leaves are close to a leaf (distance 1 or 2). In this section, we show that at large times, there are exceptional vertices that are much further away from the nearest leaf, namely at distance  $\Theta(\log n / \log \log n)$ . We recall that

$$M_n = \max_{i \leq n} \min_{\ell: D_n(\ell)=1} d_n(i, \ell)$$

is the maximal distance of any vertex to the closest leaf at time  $n$ , which we refer to as the *leaf-depth* at time  $n$ . To prove Theorem 4.3, we first need the following lemma to transfer upper bounds on the leaf-depth in the uniform random recursive trees to upper bounds on the leaf-depth in random friend trees.

**Lemma 6.4.** The coupling defined in Lemma 6.3 between the random friend tree  $(T_n, n \geq 1)$  and the uniform random recursive tree  $(T'_n, n \geq 1)$  also satisfies that for any leaf  $\ell'$  in  $T'_n$ , in  $T_n$ , the vertex  $\ell'$  is at distance at most 1 from a leaf.

*Proof.* The statement clearly holds for  $n \leq 2$ . Next, suppose that for some  $m$  the statement is satisfied in  $T_{m-1}$  and  $T'_{m-1}$ . Fix an  $\ell' \leq m$  so that  $\ell'$  is a leaf in  $T'_m$ . We claim that  $\ell'$  is at distance at most 1 from a leaf in  $T_m$ . First, if  $\ell' = m$ , then  $\ell'$  is also a leaf in  $T_m$  and the claim follows. If  $\ell' \leq m-1$ , then  $\ell'$  is also a leaf in  $T'_{m-1}$ , so by the induction hypothesis,  $\ell'$  is at distance at most 1 from a leaf  $\ell$  in  $T_{m-1}$ . If  $\ell$  is also a leaf in  $T_m$ , the claim follows. Otherwise, for  $\ell$  to be a leaf in  $T_{m-1}$ , but not  $T_m$  it is necessary that  $W_m = \ell$ . If  $W_m = \ell$  and  $d_{m-1}(\ell, \ell') = 1$ , then  $\ell'$  is the unique neighbour of  $\ell$  in  $T_{m-1}$  (because  $\ell$  is a leaf in  $T_{m-1}$ ), so  $V_m = \ell'$ , contradicting that  $\ell'$  is a leaf in  $T'_m$ . Thus, if  $W_m = \ell$ , then  $d_{m-1}(\ell, \ell') = 0$ , meaning that  $\ell = \ell'$ . Therefore, in  $T_m$ , vertex  $m$  is a leaf that is at distance 1 from  $\ell'$ .  $\square$

The next lemma gives an upper bound on the leaf-depth in the uniform random recursive tree. Together with Lemma 6.4 we deduce an upper bound for the leaf-depth in random friend trees, given in Proposition 6.6.

**Lemma 6.5.** Let  $M'_n$  be the leaf-depth in the uniform random recursive tree at time  $n$ . Then, for any  $\varepsilon > 0$ ,

$$\mathbf{P} \left\{ M'_n \geq (1 + \varepsilon) \frac{\log n}{\log \log n} \right\} = o(1).$$

*Proof.* Let  $T'_n$  be the uniform random recursive tree at time  $n$ . For any vertex  $v \in T'_n$ , we define a canonical path  $\mathcal{P}_n(v)$  in  $T'_n$  that ends in a leaf. If  $v$  is a leaf in  $T'_n$ , set  $\mathcal{P}_n(v) = v$ . Otherwise, let  $w = w_n(v)$  be the neighbour of  $v$  in  $T'_n$  with the largest label and let  $\mathcal{P}_n(v)$  be  $v$  concatenated with  $\mathcal{P}_n(w)$ . So, to obtain the path  $\mathcal{P}_n(v)$ , start from  $v$

and sequentially move to the largest labelled neighbour until reaching a leaf. Define  $T'_n(v)$ , as the connected component of  $v$  in  $T'_n$  if the edge between  $v$  and its parent was removed (or, equivalently,  $T'_n(v)$  is the subtree of  $T'_n$  that consists of all vertices that are connected to  $v$  by a path on which  $v$  is the lowest labelled vertex). Let  $|\mathcal{P}_n(v)|$  be the number of edges on the path. To prove the lemma, it is sufficient to show that  $\max_{v \in [n]} |\mathcal{P}_n(v)| \leq (1 + \epsilon) \log n / \log \log n$ . First, we check that, for any  $\ell \geq 1$ ,

$$\mathbf{P} \{ |T'_n(w_n(v))| \geq \ell \mid |T'_n(v)| = m \} = \begin{cases} \frac{1}{\ell} & \text{if } \ell < m \\ 0 & \text{otherwise,} \end{cases} \quad (6.2)$$

where  $m \leq n$ . Observe that, conditionally on  $|T'_n(v)| = m$ , if the vertices in  $T'_n(v)$  are assigned labels in  $[m]$  that respect the order of the original labels, the resulting tree has the same law as  $T'_m$ . Therefore, (6.2) follows if, for any  $m$ , we have that for all  $\ell \geq 1$ ,

$$\mathbf{P} \{ |T'_m(w_m(1))| \geq \ell \} = \begin{cases} \frac{1}{\ell} & \text{if } \ell \leq m \\ 0 & \text{otherwise,} \end{cases}$$

where we recall that  $T'_m(w_m(1))$  is the subtree rooted at the youngest child of 1.

The proof is by induction on  $m$ . The statement clearly holds for  $m = 1$ . Now, suppose the statement holds for  $m = k - 1$ . For  $m = k$ , the statement is obvious for  $\ell = 1$  and  $\ell > k$  since  $1 \leq |T'_k(w_k(1))| \leq k$ . Observe that, if  $V_k = 1$  (i.e, if vertex  $k$  connects to vertex 1), then  $T'_k(w_k(1))$  consists only of the vertex  $k$  in which case  $|T'_k(w_k(1))| = 1$ . If  $V_k \in T'_{k-1}(w_{k-1}(1))$ , then  $T'_k(w_k(1))$  is composed of the vertices of  $T'_{k-1}(w_{k-1}(1))$  and vertex  $k$ , so  $|T'_k(w_k(1))| = |T'_{k-1}(w_{k-1}(1))| + 1$ . If  $V_k \notin \{1\} \cup T'_{k-1}(w_{k-1}(1))$  then  $T'_k(w_k(1)) = T'_{k-1}(w_{k-1}(1))$ , giving  $|T'_k(w_k(1))| = |T'_{k-1}(w_{k-1}(1))|$ . Therefore, for  $1 < \ell \leq k$ ,

$$\begin{aligned} \{ |T'_k(w_k(1))| \geq \ell \} &= \{ \{ |T'_{k-1}(w_{k-1}(1))| \geq \ell \} \cap \{ V_k \neq 1 \} \} \\ &\cup \{ \{ |T'_{k-1}(w_{k-1}(1))| = \ell - 1 \} \cap \{ V_k \in T'_{k-1}(w_{k-1}(1)) \} \}. \end{aligned}$$

By the induction hypothesis, for  $\ell < k$ ,

$$\mathbf{P} \{ |T'_k(w_k(1))| \geq \ell \} = \frac{1}{\ell} \frac{k-1}{k} + \frac{1}{\ell(\ell-1)} \frac{\ell-1}{k} = \frac{1}{\ell},$$

and for  $\ell = k$ ,

$$\mathbf{P} \{ |T'_k(w_k(1))| \geq k \} = 0 + \frac{1}{k-1} \frac{k-1}{k} = \frac{1}{k}.$$

The equality (6.2) follows for all  $m$ . Now, define  $w^{(1)} = w_n(v)$  the vertex at distance 1 from  $v$  on  $\mathcal{P}_n(v)$  and  $w^{(\ell)} = w_n(w^{(\ell-1)})$  the vertex at distance  $\ell$  from  $v$  on  $\mathcal{P}_n(v)$ . Then,

$$\{ |\mathcal{P}_n(v)| \geq k \} = \{ |T'_n(w^{(1)})| \geq k \} \cap \{ |T'_n(w^{(2)})| \geq k-1 \} \cap \dots \cap \{ |T'_n(w^{(k)})| \geq 1 \}.$$

Together with (6.2), this implies that

$$\mathbf{P} \{ |\mathcal{P}_n(v)| \geq k \} \leq \frac{1}{k!} \leq \frac{e^k}{k^k}.$$

Then, fix  $\epsilon > 0$  and substitute  $(1 + \epsilon) \frac{\log n}{\log \log n}$  to  $k$ . The above equation directly implies that

$$\mathbf{P} \left\{ |\mathcal{P}_n(v)| \geq (1 + \epsilon) \frac{\log n}{\log \log n} \right\} = o(n^{-1}).$$

Finally, a union bound implies Lemma 6.5.  $\square$

From Lemmas 6.4 and 6.5 we obtain an upper bound on the leaf depth in random friend trees.

**Proposition 6.6.** For any  $\varepsilon > 0$

$$\mathbf{P} \left\{ M_n \geq (2 + \varepsilon) \frac{\log n}{\log \log n} \right\} = o(1).$$

*Proof.* Couple the random friend tree  $(T_n, n \geq 1)$  and the uniform random recursive tree  $(T'_n, n \geq 1)$  as in Lemma 6.4. Then, fix  $i \leq n$  such that for  $N'_n(i)$ , the degree of vertex  $i$  in  $T'_n$ , we see that

$$\begin{aligned} \min_{\ell: D_n(\ell)=1} d_n(i, \ell) &\leq \min_{\ell: D_n(\ell)=1} \min_{\ell': N'_n(\ell')=1} (d_n(i, \ell') + d_n(\ell', \ell)) \\ &\leq \min_{\ell': N'_n(\ell')=1} d_n(i, \ell') + 1 \\ &\leq \min_{\ell': N'_n(\ell')=1} 2d'_n(i, \ell') + 1, \end{aligned}$$

where the last two inequalities follow from the properties of the coupling. By taking the maximum over  $i \in [n]$ , we have  $M_n \leq 2M'_n + 1$  and so Lemma 6.5 implies the proposition.  $\square$

To conclude the proof of Theorem 4.3 it remains to prove an asymptotic lower bound for  $M_n$ . In order to show that the leaf depth is at least of order  $\log n / \log \log n$  we first present a proof of the corresponding result for the URRT. To the best of our knowledge, this result does not appear elsewhere. Moreover, the proof is less technical but has the same structure as its counterpart for random friend trees, so it is a good way to introduce the ideas needed for our main proof. In doing so, we hope that the technicalities in the proof of Lemma 6.8 are easier to understand.

**Proposition 6.7.** Let  $M'_n$  be the leaf-depth in the uniform random recursive tree at time  $n$ . Then,  $M'_n = \Omega_p(\log n / \log \log n)$ .

*Proof.* Let  $T'_n$  be the uniform random recursive tree at time  $n$ . In a uniform random recursive tree, the number of vertices with degree at least 3 goes to infinity almost surely (see Janson [11]). We call such a vertex a branch point. Therefore, we can choose  $n$  sufficiently large such that there is at least one branch point. Let  $P'_n$  be the maximal distance of any leaf in  $T'_n$  to the nearest branch point, that is,

$$P'_n = \max_{\ell: N'_n(\ell)=1} \min_{j: N'_n(j) \geq 3} d'_n(\ell, j).$$

We remark that  $M'_n \geq P'_n/2$  because a midpoint of a longest leaf-to-branchpoint path is at distance at least  $P'_n/2$  from the nearest leaf. It is well known (see Janson [11]) that the proportion of leaves in an URRT tends to  $1/2$ , consequently  $X'_m \geq 2/m \rightarrow 1/2$  almost surely. Thus, for  $\varepsilon > 0$  and  $n$  sufficiently large,

$$\mathbf{P} \left\{ 1 - \frac{X'_k \geq 2}{k} \leq \frac{1}{3}; \forall k \geq n \right\} > 1 - \varepsilon.$$

Now, condition on the number of leaves at time  $n/2$  being at least  $n/6$ , that is  $X'_{n/2} \geq n/6$ , and let  $\{v_1, \dots, v_P\}$  be an arbitrary set of  $P = n/6$  leaves of  $T'_{n/2}$ . We study the subtrees rooted at  $v_i$  and will show that at time  $n$ , with high probability, for at least one  $i \in [P]$  the subtree of  $v_i$  contains a path from  $v_i$  ending in a leaf and solely consisting of  $\Omega(\log n / \log \log n)$  vertices of degree two. To this end, we say that a path consisting of degree two vertices that ends in a leaf *grows* at time  $m$  if vertex  $m + 1$  attaches to the

leaf at the end of the path. This increases the length of the path by 1. We say that a path consisting of degree two vertices that ends in a leaf *dies* at time  $m$  if vertex  $m + 1$  connects to a degree-two vertex on the path that is at distance at most  $K$  from the leaf. That is, we only keep track of paths up to distance  $K$  from the leaf. Note that at each time step only one path can grow or die. We will only track paths until their first death. Note that, at time  $m$ , for  $i \in [P]$ , conditionally on the path rooted at  $v_i$  has not died yet, the probability of the path growing is  $1/m$  and the probability of the path dying is at most  $K/m$ . Observe that  $P'_n$  stochastically dominates the minimum of  $K$  and the length of the longest path at time  $n$  rooted at some  $v_i$  for  $i \in [P]$  that has not died. We can then couple these path growth processes to a balls-in-bins model as follows. Let  $P$  be the number of bins. The process is started at time  $n/2$  with  $P$  empty bins. For each time  $m \in [n/2, n]$ , with probability  $P \cdot K/m$  add a black ball to a uniform random bin, or with probability  $P/m$  add a white ball to a uniform random bin. By doing so, a black ball is added to a given bin with probability  $K/m$  and a white ball is added to a given bin with probability  $1/m$ . We further see that  $P'_n$  stochastically dominates the smallest number between  $K$  and the maximum number of white balls at time  $n$  in a bin with zero black balls. It therefore suffices to show that for some  $c > 0$  and  $K := c \log n / \log \log n$ , for  $n$  large enough, with high probability, one of the  $P$  bins contains at least  $K$  white balls and no black balls. Observe that, for  $n$  sufficiently large, between times  $n/2$  and  $n$ , with probability at least  $1 - \epsilon$ , at most  $B = 2PK$  black balls and at least  $W = P/8$  white balls are added. Conditioned on this event, the probability that a specific bin contains at least  $K$  white balls is at least

$$\binom{W}{K} P^{-K} (1 - P^{-1})^{W-K} > e^{-1} \left( \frac{W}{PK} \right)^K = e^{-1} \left( \frac{1}{8K} \right)^K,$$

for large  $n$ , where we bound  $(1 - P^{-1})^{W-K} \geq (1 - P^{-1})^W = ((1 - 6/n)^{n/6})^{1/8} > e^{-1}$ . Therefore the expected number of bins containing at least  $K$  white balls can be bounded from below by

$$P \cdot e^{-1} (8K)^{-K} = e^{-1} \frac{n}{6} \left( \frac{\log \log n}{8c \log n} \right)^{c \log n / \log \log n} > n^{1-c}$$

for large  $n$ .

Then, if  $c < 1/2$ , since the numbers of white balls in two distinct bins have negative covariance, Chebyshev's inequality bounds the probability that the number of paths with at least  $K$  growth events is less than  $n^{1-2c}$  tends to 0, so in particular is smaller than  $\epsilon$  for  $n$  sufficiently large. Any of these has 0 black balls with probability at least  $(1 - P^{-1})^B \geq e^{-3K} = \omega(n^{1-2c})$ , so another straightforward application of the second moment method implies that for  $n$  large enough, for any  $c < 1/2$ , with probability at least  $1 - 3\epsilon$  at time  $n$ , there is a bin containing  $c \log(n) / \log \log(n)$  white balls and no black balls. Recalling the stochastic dominance, that is valid on an event of probability at least  $1 - \epsilon$ , with probability at least  $1 - 4\epsilon$  there is a leaf at distance at least  $c \log n / \log \log n$  from the nearest branch point.  $\square$

Using a similar proof, we can prove an asymptotic lower bound for the leaf-depth in random friend trees.

**Proposition 6.8.** Let  $M_n$  be the leaf-depth in the random friend tree at time  $n$ . Then,  $M_n = \Omega(\log n / \log \log n)$  in probability.

*Proof.* By Theorem 4.7,  $X_n^{\geq 3} \rightarrow \infty$  almost surely, so we may pick  $n$  large enough such that  $X_n^{\geq 3} \geq 1$ . Then, define

$$P_n = \max_{\ell: D_n(\ell)=1} \min_{j: D_n(j) \geq 3} d_n(\ell, j)$$

to be the largest distance from a leaf to the nearest branch point in  $T_n$ . We see that  $M_n \geq P_n/2$  because the midpoint of the longest leaf-to-branchpoint path is at distance at least  $P_n/2$  from the nearest leaf. Therefore, it suffices to show that  $P_n = \Omega(\log n / \log \log n)$  in probability.

As in Proposition 6.7 the proof is divided into two steps. We first show that, for  $n$  large enough, with high probability at time  $n/2$  there are at least  $n^{\delta/2}$  leaves attached to a degree two vertex. Remark that this step requires slightly more work than the URRT case. For the URRT, one only needs to ensure the presence of many leaves at time  $n/2$  to guarantee that that a long path of subsequent degree-two vertex will emerge from one of these leaves. For the random friend tree, one needs to ensure that, at time  $n/2$ , there exists many leaves attached to a degree two vertex, to guarantee that new vertices attach to these leaves with some sufficient probability. This difference is due to the attachment rule in random friend trees. Denote by  $\mathcal{L}_n$  the set of leaves attached to a vertex of degree two at time  $n$  and let  $H_n = |\mathcal{L}_n|$  be the number of leaves attached to a degree two vertex in  $T_n$ . In the second part of the proof, we show that it is likely that at time  $n$ , a path of subsequent degree two vertices longer than  $c \log n / \log \log n$  grows from at least one of the leaves of  $\mathcal{L}_{n/2}$ , that is,

$$\max_{\ell \in \mathcal{L}_n} \min_{j: D_n(j) \geq 3} d_n(\ell, j) > c \frac{\log n}{\log \log n},$$

with high probability. Lemma 6.15 states that

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \geq \frac{1}{3n} X_n^{\geq 3},$$

and by Theorem 4.7,  $n^{-\delta} X_n^{\geq 3} \rightarrow \infty$  almost surely, for  $0.1 < \delta < 0.9$ . Therefore, for fixed  $\epsilon > 0$  and sufficiently large  $n$ ,

$$\mathbf{P} \left\{ \frac{1}{3m} X_m^{\geq 3} > m^{-(1-\delta)} \forall m \geq n/4 \right\} \geq 1 - \epsilon.$$

It follows that, for all  $n/4 \leq m \leq n/2$ ,

$$\mathbf{P} \left\{ \Delta X_m^{\geq 2} = 1 \mid \frac{1}{3m} X_m^{\geq 3} > m^{-(1-\delta)} \right\} \geq n^{-(1-\delta)}$$

because  $m^{-(1-\delta)} \geq n^{-(1-\delta)}$  if  $m \leq n/2$ .

We will show that  $H_n$  grows polynomially in probability. Note that for every vertex  $v \in \mathcal{L}_m$ ,  $\mathbf{P} \{W_m = v\} = 1/(2m)$  and  $\Delta H_m = 1$  if and only if  $W_m$  is a leaf in  $T_m$  that is not in  $\mathcal{L}_m$ . Further,  $\Delta X_m^{\geq 2} = 1$  if and only if  $W_m$  is a leaf in  $T_m$ . By summing over all possible values of  $W_m$  we therefore get the lower bound

$$\begin{aligned} & \mathbf{P} \left\{ \Delta H_m = 1 \mid H_m = k, \frac{1}{3m} X_m^{\geq 3} > m^{\delta-1} \right\} \\ & \geq \mathbf{P} \left\{ \Delta X_m^{\geq 2} = 1 \mid \frac{1}{3m} X_m^{\geq 3} > m^{-(1-\delta)} \right\} - \frac{k}{2m} \geq n^{\delta-1} - \frac{2k}{n}, \end{aligned}$$

for  $n/4 \leq m \leq n/2$ . Therefore, either  $H_m \geq n^{\delta}/16$  or  $\mathbf{P} \{ \Delta H_m = 1 \mid \frac{1}{3m} X_m^{\geq 3} > m^{-(1-\delta)} \} \geq \frac{7}{8} n^{-(1-\delta)}$  for  $n/4 \leq m \leq n/2$ .

Finally,  $\Delta H_m = -1$  if  $W_m$  is a degree two vertex attached to a leaf, and there are exactly  $H_m$  such vertices in  $T_m$ . For  $v$  a given vertex of degree two attached to a leaf,  $\mathbf{P} \{W_m = v\} \leq 2/m \leq 8/n$  for  $n/4 \leq m \leq n/2$ . By a union bound, for  $n/4 \leq m \leq n/2$ ,

$$\mathbf{P} \{ \Delta H_m = -1 \mid H_m = k \} \leq 8k/n.$$

By the two arguments above, for  $n/4 \leq m \leq n/2$ ,

$$\mathbf{P} \left\{ \Delta H_m = -1 \mid H_m \leq \frac{1}{16} n^\delta \right\} \leq \frac{1}{2} n^{-(1-\delta)},$$

and

$$\mathbf{P} \left\{ \Delta H_m = 1 \mid H_m \leq \frac{1}{16} n^\delta, \frac{1}{3m} X_m^{\geq 3} > m^{\delta-1} \right\} \geq \frac{7}{8} n^{-(1-\delta)},$$

Then, it follows that, on the event  $\{\forall m \in [n/4, n/2], \frac{1}{3m} X_m^{\geq 3} > m^{\delta-1}\}$ , with probability at least  $1 - \varepsilon$ , for  $n$  sufficiently large,  $H_m$  grows to at least  $\frac{1}{17} n^\delta$  between times  $n/4$  and  $n/2$ . Finally, for  $n$  large enough,  $n^\delta/17 > n^{\delta/2}$ , so

$$\mathbf{P} \left\{ H_{n/2} \geq n^{\delta/2} \right\} \geq 1 - 2\varepsilon,$$

which concludes the first part of the proof.

We now condition on the event  $\{H_{n/2} \geq n^{\delta/2}\}$ . We show that, on this event, it is likely that at time  $n$  at least one of the leaves of  $\mathcal{L}_{n/2}$  is at distance at least  $\log n / \log \log n$  from the nearest branch point. The second part of the proof is identical from the URRT proof, but we present it again for clarity.

Let  $\{v_1, \dots, v_P\}$  be an arbitrary set of  $P = n^{\delta/2}$  leaves in  $\mathcal{L}_{n/2}$ . We study the subtrees rooted at  $v_i$  and will show that at time  $n$ , with high probability, for at least one  $i \in [P]$  the subtree of  $v_i$  contains a path ending in a leaf and solely consisting of  $\Omega(\log n / \log \log n)$  vertices of degree two. To this end, we say that a path consisting of degree two vertices that ends in a leaf *grows* at time  $m$  if vertex  $m + 1$  attaches to the leaf at the end of the path. This increases the length of the path by 1. We say that a path consisting of degree two vertices that ends in a leaf *dies* at time  $m$  if vertex  $m + 1$  connects to a degree two vertex on the path that is at distance at most  $K$  from the leaf. That is, we only keep track of paths of length at most  $K$ . Note that, at each time step, only one path can grow or die. We will only track paths until their first death. Also note that, at time  $m \in [n/2, n]$ , for  $i \in [P]$ , conditionally on the path rooted at  $v_i$  has not died yet, the probability of the path growing is  $1/(2m)$  and the probability of the path dying is at most  $K/m$ . Observe that  $P_n$  stochastically dominates the minimum of  $K$  and the longest path at time  $n$  rooted at some  $v_i$  for  $i \in [P]$  that has not died. We can then couple these path growth processes to a balls-in-bins model. Let  $P$  be the number of bins and start the process with  $P$  empty bins. For each time  $m \in [n/2, n]$ , with probability  $P \cdot K/m$  add a black ball to a uniform random bin, or with probability  $P/(2m)$  add a white ball to a uniform random bin. By doing so, a black ball is added to a given bin with probability  $K/m$  and a white ball is added to a given bin with probability  $1/(2m)$ . We further see that  $P_n$  stochastically dominates the smallest number between  $K$  and the maximum number of white balls at time  $n$  in a bin with zero black balls. It therefore suffices to show that, for some  $c > 0$  and  $K := c \log n / \log \log n$ , with high probability at least one of the  $P$  bins contains at least  $K$  white balls and no black balls. Observe that, for  $n$  sufficiently large, between time  $n/2$  and  $n$ , with probability at least  $1 - \varepsilon$ , at most  $B = 2PK$  black balls and at least  $W = P/8$  white balls are added. Conditioned on this event, the probability that a specific bin has at least  $K$  white balls is at least

$$\binom{W}{K} P^{-K} (1 - P^{-1})^{W-K} > e^{-1} \left( \frac{W}{PK} \right)^K = e^{-1} \left( \frac{1}{8K} \right)^K,$$

for large  $n$ , where we bound  $(1 - P^{-1})^{W-K} \geq (1 - P^{-1})^W = ((1 - n^{-\delta/2})^{n^{-\delta/2}})^{1/8} > e^{-1}$ . The expected number of bins containing at least  $K$  white balls is bounded from below by

$$P \cdot e^{-1}(8K)^{-K} = e^{-1}n^{\delta/2} \left( \frac{\log \log n}{c \log n} \right)^{c \log n / \log \log n} > n^{\delta/2-c},$$

for large  $n$ .

Then, if  $c < \delta/8$ , since the numbers of white balls in two distinct bins have negative covariance, Chebyshev's association inequality gives that the probability that the number of paths with at least  $K$  growth events is less than  $n^{\delta/2-2c}$  tends to 0, so in particular, is smaller than  $\epsilon$  for  $n$  sufficiently large. Any of these contains no black balls with probability at least  $(1 - P^{-1})^B \geq e^{-3K} = \omega(n^{\delta/2-2c})$ , so another straightforward application of the second moment method implies that for  $n$  large enough, for any  $c < \delta/8$ , with probability at least  $1 - 3\epsilon$  there is a bin with  $c \log(n)/\log \log(n)$  white balls and no black balls at time  $n$ . Recalling the statistical dominance, that is valid on an event of probability at least  $1 - 2\epsilon$ , with probability at least  $1 - 5\epsilon$ , there is a leaf at distance at least  $c \log n / \log \log n$  from the nearest branch point.  $\square$

#### 6.4 High-degree vertices

[Proof of Theorems 4.5 and 4.6]

Recall that  $Z_v = \liminf_{n \rightarrow \infty} \frac{D_n(v)}{n}$ , and that by Theorem 3.1, in fact

$$Z_v = \lim_{n \rightarrow \infty} \frac{D_n(v)}{n} \text{ almost surely.}$$

The following lemma, combined with Theorem 4.7, gives a lower bound on the number of hubs and proves Theorem 4.5.

**Lemma 6.9.** Almost surely,

$$\#\{v \in [n] : Z_v > 0\} > \frac{1}{2} X_n^{\geq 2}.$$

We prove this lemma using Lemma 6.10, below, but we need some additional definitions for its statement. For a graph  $G = (V, E)$ , we say  $V' \subset V$  is an *edge cover* of  $G$  if for each  $e \in E$ , there is a  $v \in V'$  such that  $v \in e$ . Define the *minimal edge cover number* of a graph  $G = (V, E)$ , denoted by  $EC(G)$ , as follows

$$EC(G) := \min\{|V'| : V' \text{ is an edge cover of } G\}. \quad (6.3)$$

Lemma 6.9 is a direct consequence of Theorem 3.2, which states that each edge contains a hub, and the following lemma.

**Lemma 6.10.** For any tree  $t$ , we have that

$$|EC(t)| \geq \frac{1}{2} X^{\geq 2}(t),$$

where  $X^{\geq 2}(t)$  denotes the number of non-leaves in the tree  $t$ .

*Proof.* We can assume that  $t$  is a rooted tree by declaring an arbitrary vertex in  $t$  the root. Decompose the tree  $t$  into the following vertex-disjoint paths. Let  $\ell_1, \dots, \ell_{X^1(t)}$  be the leaves of  $t$ . For a leaf  $\ell$  of  $t$ , let  $P(\ell)$  be the path from  $\ell$  to the root of  $t$ . For each  $i \in [X^1(t)]$ , let  $P'(\ell_i) = P(\ell_i) \setminus \bigcup_{j=0}^{i-1} P'(\ell_j)$ , that is,  $P'(\ell_i)$  is the path  $P(\ell_i)$  stripped of the vertices in  $\bigcup_{j=0}^{i-1} P'(\ell_j)$ . This decomposition gives us  $X^1(t)$  disjoint paths  $P'_{\ell_1}, \dots, P'_{\ell_{X^1(t)}}$ . Note that if  $V' \subset V$  is an edge cover of  $t$  it must also be an edge cover of  $P'(\ell_1) \cup \dots \cup P'(\ell_{X^1(t)})$  (indeed, while removing edges, the requirement for a collection of edges to be an edge cover is weakened). An edge cover of a disconnected graph is a disjoint union of edge



covers of the components, and an edge cover of a path of  $m$  vertices contains at least  $\lfloor m/2 \rfloor$  vertices, so

$$|EC(t)| \geq \sum_{i=1}^{X^1(t)} \left\lfloor \frac{|P'_i|}{2} \right\rfloor \geq \sum_{i=1}^{X^1(t)} \left( \frac{|P'_i| - 1}{2} \right) = \frac{|V(t)| - X^1(t)}{2} = \frac{X^{\geq 2}(t)}{2}. \quad \square$$

We now prove Theorem 4.6, which in particular implies that for any  $k$ , the number of vertices with degree at least  $k$  goes to infinity almost surely.

*Proof of Theorem 4.6.* Fix a constant  $M \in \mathbb{N}$  and let  $(m_n, n \geq 1)$  be a sequence satisfying  $m_n = o(n)$ . We will prove that  $\liminf_{n \rightarrow \infty} X_n^{\geq m_n} \geq M$  almost surely. By Lemma 6.1, there exists an almost surely finite time  $\tau$  such that the diameter of the tree at time  $\tau$  exceeds  $2M$ . Fix an arbitrary path of  $2M+1$  vertices in  $T_\tau$  and, for  $n \geq \tau$ , let  $N_n^{(1)} \geq \dots \geq N_n^{(2M+1)}$  be the degrees of the vertices on this path in decreasing order. Then, by Theorem 3.2, almost surely, at least  $M$  vertices on this path are hubs, so

$$\liminf_{n \rightarrow \infty} \frac{N_n^{(M)}}{n} > 0$$

and in particular, there is a finite time  $\tau'$  such that  $N_n^{(M)} > m_n$  for all  $n \geq \tau'$ . This implies that from time  $\tau'$  onwards, there are at least  $M$  vertices with degree at least  $m_n$ , so

$$\mathbf{P} \left\{ \liminf_{n \rightarrow \infty} X_n^{\geq m_n} \geq M \right\} = 1. \quad \square$$

## 6.5 Low-degree vertices

[Proof of Theorems 4.7 and 4.8] In this subsection we prove polynomial upper and lower bounds of  $X_n^{\geq k}$ , for  $k$  bounded and show that  $X_n^{\geq k} = \Theta(X_n^{\geq 2})$  almost surely. We begin by stating a general result on adapted processes, which will be of use in the proof of Theorem 4.7. Its proof can be found in the appendix. For each  $k \in \mathbb{N}$ , recall that  $X_n^k$  is the number of vertices of degree  $k$  in  $T_n$  and  $X_n^{\geq k}$  is the number of vertices of degree at least  $k$  in  $T_n$ .

**Proposition 6.11.** Let  $(X_k)_{k \geq 0}$  and  $(Y_k)_{k \geq 0}$  be integer-valued non-decreasing processes adapted to some filtration  $(\mathcal{F}_k)_{k \geq 0}$  such that  $0 \leq \Delta X_k + \Delta Y_k \leq 1$  for all  $k$ . Suppose there exists  $\alpha > 0$  such that  $\mathbf{E}[\Delta X_k | \mathcal{F}_k] \geq \alpha \mathbf{E}[\Delta Y_k | \mathcal{F}_k]$  on the event  $\{X_k < \alpha Y_k\}$ , except at finitely many times almost surely. Then for any  $\beta \in (0, \alpha)$  there exists  $C > 0$  such that

$$\mathbf{P} \{X_n < \beta Y_n - C \log n \text{ infinitely often}\} = 0,$$

and for all  $n$ ,  $\mathbf{E}[X_n] \geq \beta \mathbf{E}[Y_n] - C \log n$ .

**Lemma 6.12.** For  $n \geq 4$ ,

$$\mathbf{E}[\Delta X_n^{\geq 2} | T_n] \leq \frac{1}{2n} X_n^{\geq 2} + \frac{1}{2n} X_n^{\geq 3}, \quad (6.4)$$

and for  $n \geq 5$ ,

$$\mathbf{E}[\Delta X_n^{\geq 3} | T_n] \leq \frac{4}{3n} X_n^{\geq 3}. \quad (6.5)$$

*Proof.* To prove (6.4), note that  $\Delta X_n^{\geq 2} > 0$  precisely if  $W_n$  is a leaf of  $T_n$ , so

$$\mathbf{E}[\Delta X_n^{\geq 2} | T_n] = \frac{1}{n} \sum_{v \in T_n} \frac{L_n(v)}{D_n(v)}.$$

When  $n \geq 3$ , leaves do not have leaves as neighbours, and when  $n \geq 4$ , any vertex  $v$  of degree two in  $T_n$  has at most one leaf neighbour, thus if  $v$  has degree two,  $L_n(v)/D_n(v) = L_n(v)/2 \leq 1/2$ . Together with the previous equality, this implies that

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \leq \frac{1}{n} \left( \sum_{\{v: D_n(v)=2\}} \frac{1}{2} + \sum_{\{v: D_n(v) \geq 3\}} \frac{L_n(v)}{D_n(v)} \right) \leq \frac{X_n^2}{2n} + \frac{X_n^{\geq 3}}{n}.$$

Since  $X_n^2 = X_n^{\geq 2} - X_n^{\geq 3}$ , this implies (6.4).

To prove (6.5), for  $1 \leq i \leq j$  let

$$S_{ij} = \{v \in T_n : D_n(v) = 2, \text{ the neighbours of } v \text{ in } T_n \text{ have degrees } i \text{ and } j\}.$$

Then  $X_n^2 = \sum_{1 \leq i \leq j} |S_{ij}|$  and

$$\mathbf{E} [\Delta X_n^{\geq 3} \mid T_n] = \sum_{1 \leq i \leq j} \mathbf{P} \{W_n \in S_{ij} \mid T_n\} = \frac{1}{n} \sum_{1 \leq i \leq j} |S_{ij}| \left( \frac{1}{i} + \frac{1}{j} \right).$$

We bound this sum by splitting it into three sums. First, for terms with  $i = 1$  and  $j \geq 3$  we have  $1/i + 1/j \leq 4/3$ . For  $v \in \bigcup_{2 \leq i \leq j} S_{ij}$  we have  $1/i + 1/j \leq 1$ , while for  $v \in S_{12}$  we have  $1/i + 1/j = 3/2$ . We claim that at most half of the vertices in  $S_{12} \cup \bigcup_{2 \leq i \leq j} S_{ij}$  can be in  $S_{12}$ . Indeed, provided that  $n \geq 5$ , if  $v \in S_{12}$  then for  $u$  its unique neighbour with degree 2 it holds that  $u \in \bigcup_{2 \leq i \leq j} S_{ij}$ . Moreover,  $n \geq 5$  implies that  $u$  has at most one neighbour in  $S_{12}$ . Therefore,  $|S_{12}| \leq |\bigcup_{2 \leq i \leq j} S_{ij}|$ . This implies that

$$\sum_{1 \leq i \leq j} |S_{ij}| \left( \frac{1}{i} + \frac{1}{j} \right) \leq \frac{4}{3} \sum_{j \geq 3} |S_{1j}| + \frac{1}{2} \left( 1 + \frac{3}{2} \right) \left( |S_{12}| + \sum_{2 \leq i \leq j} |S_{ij}| \right) \leq \frac{4}{3} X_n^2,$$

therefore  $\mathbf{E} [\Delta X_n^{\geq 3} \mid T_n] \leq \frac{4}{3n} X_n^2$  as claimed.  $\square$

Before stating the next lemma we introduce the notation  $X_n^{k, \leq k}$ , the number of vertices of degree  $k$  having at most one neighbour of degree at least  $k+1$ , and  $X_n^{k, > k}$ , the number of vertices of degree  $k$  with at least two neighbours of degree at least  $k+1$ .

**Lemma 6.13.** For any positive integer  $k$ ,  $n \geq 3$ ,

$$X_n^k = X_n^{k, > k} + X_n^{k, \leq k}, \quad (6.6)$$

$$X_n^{\geq k+1} \geq X_n^{k, > k}, \quad (6.7)$$

$$\mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] \geq \frac{k-1}{kn} X_n^{k, \leq k}. \quad (6.8)$$

*Proof.* The equality (6.6) follows directly from the definition of  $X_n^{k, > k}$  and  $X_n^{k, \leq k}$ . To prove the second statement, remark that any vertex  $v$  contributing to  $X_n^{k, > k}$  has at least two neighbours of degree at least  $k+1$ , and at least one is a child of  $v$ . Since every vertex is the child of at most 1 vertex, (6.7) follows.

Finally, to prove (6.8), one must understand how vertices of degree  $k+1$  are created. In order to have  $\Delta X_n^{\geq k+1} = 1$ , it is sufficient (but not necessary) that vertex  $n+1$  attaches to a vertex counted by  $X_n^{k, \leq k}$ , or in other words, that  $W_n = w$  for some vertex  $w$  with degree  $k$  which has at least  $k-1$  neighbours of degree at most  $k$ . For each such vertex  $w$ , this happens with probability at least  $\frac{k-1}{kn}$ . This proves (6.8).  $\square$

**Lemma 6.14.** For any integer  $k$ , whenever  $n \geq k+2$ .

$$\mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] \leq \frac{k-1/2}{n} X_n^k. \quad (6.9)$$

*Proof.* A vertex of degree  $k + 1$  is created at time  $n$  if  $W_n$  has degree  $k$  in  $T_n$ . For a vertex  $w \in T_n$  of degree  $k$ , the probability of  $W_n = w$  is maximized if the neighbours of  $w$  have lowest possible degree; that is, if  $w$  has  $k - 1$  leaf neighbours and one neighbour of degree two. In this case, the probability that  $W_n = w$  equals  $\frac{k-1/2}{n}$ ; the lemma follows.  $\square$

We use the following two lemmas to show that the number of non-leaves grows at least polynomially.

**Lemma 6.15.** For  $n \geq 3$ ,

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \geq \frac{1}{3n} X_n^{\geq 3}.$$

*Proof.* Note that  $\Delta X_n^{\geq 2} = 1$  if and only if  $W_n$  is a leaf. Therefore, as observed before, if  $n \geq 3$  then

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] = \frac{1}{n} \sum_{\{v: D_n(v) \geq 2\}} \frac{L_n(v)}{D_n(v)}. \quad (6.10)$$

Let  $T'_n$  be equal to  $T_n$  with all of its leaves removed. We consider different sets of vertices in  $T'_n$  and we study their contribution to the sum above.

Let  $V_1$  be the set of vertices of  $T_n$  that are leaves in  $T'_n$  and that were vertices of degree 2 in  $T_n$ , so that for each  $v \in V_1$ ,  $\frac{L_n(v)}{D_n(v)} = \frac{1}{2}$ . Let  $V_2$  be the vertices of  $T_n$  that are leaves in  $T'_n$  that were vertices of degree at least 3 in  $T_n$  so that for each  $v \in V_2$ ,  $\frac{L_n(v)}{D_n(v)} = \frac{D_n(v)-1}{D_n(v)} \geq \frac{2}{3}$ . Let  $V_3$  be the vertices that have degree 2 in  $T'_n$  and had degree at least 3 in  $T_n$ , so that for each  $v \in V_3$ ,  $\frac{L_n(v)}{D_n(v)} = \frac{D_n(v)-2}{D_n(v)} \geq \frac{1}{3}$ . Finally, let  $V_4$  be the vertices that have degree at least 3 in  $T'_n$ . Therefore,

$$\sum_{\{v: D_n(v) \geq 2\}} \frac{L_n(v)}{D_n(v)} \geq \frac{1}{2}|V_1| + \frac{2}{3}|V_2| + \frac{1}{3}|V_3|. \quad (6.11)$$

To lower bound this sum note that  $|V_1| + |V_2|$  is the number of leaves in  $T'_n$ . Since  $T'_n$  is a tree, the number of leaves in  $T'_n$  is given by

$$\sum_{\{v: |\mathcal{N}(v, T'_n)| \geq 3\}} (|\mathcal{N}(v, T'_n)| - 2) + 2$$

and so

$$|V_1| + |V_2| = \sum_{v \in V_4} (|\mathcal{N}(v, T'_n)| - 2) + 2.$$

Finally, since  $\sum_{v \in V_4} (|\mathcal{N}(v, T'_n)| - 2) + 2 \geq |V_4|$  we obtain that

$$|V_1| + |V_2| \geq |V_4|,$$

hence

$$\frac{1}{2}|V_1| + \frac{2}{3}|V_2| \geq \frac{1}{3}|V_2| + \frac{1}{3}|V_4|.$$

It also holds that  $X_n^{\geq 3} = |V_2| + |V_3| + |V_4|$ , so we conclude that

$$\frac{1}{2}|V_1| + \frac{2}{3}|V_2| + \frac{1}{3}|V_3| \geq \frac{1}{3}X_n^{\geq 3}. \quad \square$$

Combined with (6.10) and (6.11), this completes the proof.

**Lemma 6.16.** Let  $\alpha = (\sqrt{13} - 3)/2 \approx 0.303$  be the unique positive solution of  $x = \frac{1-2x}{1+x}$ . Then, for any  $\beta \in (0, \alpha)$  there exists  $c > 0$  such that

$$\mathbf{P} \{X_n^{\geq 3} < \beta X_n^{\geq 2} - c \log n \text{ infinitely often}\} = 0,$$

and for all  $n$ ,  $\mathbf{E} [X_n^{\geq 3}] \geq \beta \mathbf{E} [X_n^{\geq 2}] - c \log n$ .

*Proof.* The statements follow directly by applying Proposition 6.11, once we show that for  $\alpha$  as in the Lemma statement, for  $n \geq 4$ , either  $X_n^{\geq 3} \geq \alpha X_n^{\geq 2}$  or  $\mathbf{E} [\Delta X_n^{\geq 3} | T_n] \geq \alpha \mathbf{E} [\Delta X_n^{\geq 2} | T_n]$ . Suppose that  $X_n^{\geq 3} < \alpha X_n^{\geq 2}$ . Then, by (6.4),

$$\mathbf{E} [\Delta X_n^{\geq 2} | T_n] \leq \frac{1}{2n} X_n^{\geq 2} + \frac{1}{2n} X_n^{\geq 3} \leq \frac{1+\alpha}{2n} X_n^{\geq 2}. \quad (6.12)$$

Moreover, observe that the case  $k = 2$  of Lemma 6.13 gives that

$$X_n^2 = X_n^{2,>2} + X_n^{2,\leq 2}, \quad (6.13)$$

$$X_n^{\geq 3} \geq X_n^{2,>2}, \text{ and} \quad (6.14)$$

$$\mathbf{E} [\Delta X_n^{\geq 3} | T_n] \geq \frac{1}{2n} X_n^{2,\leq 2}. \quad (6.15)$$

Since  $X_n^{\geq 3} < \alpha X_n^{\geq 2}$ , (6.14) implies that  $X_n^{2,>2} < \alpha X_n^{\geq 2}$ . Note that  $X_n^{\geq 2} = X_n^{2,>2} + X_n^{2,\leq 2} + X_n^{\geq 3}$ , so the bounds  $X_n^{2,>2} < \alpha X_n^{\geq 2}$  and  $X_n^{\geq 3} < \alpha X_n^{\geq 2}$  together imply that  $X_n^{2,\leq 2} > (1 - 2\alpha) X_n^{\geq 2}$ . Combining this bound with (6.15) and (6.12), we conclude that

$$\mathbf{E} [\Delta X_n^{\geq 3} | T_n] \geq \frac{1-2\alpha}{2n} X_n^{\geq 2} \geq \frac{1-2\alpha}{1+\alpha} \mathbf{E} [\Delta X_n^{\geq 2} | T_n] = \alpha \mathbf{E} [\Delta X_n^{\geq 2} | T_n],$$

as required.  $\square$

**Lemma 6.17.** As  $n \rightarrow \infty$ ,  $X_n^{\geq 3} \rightarrow \infty$  almost surely.

*Proof.* By Theorem 4.2, the diameter of  $T_n$  is  $\Theta(\log n)$  almost surely. Thus,

$$X_n^{\geq 2} \rightarrow \infty \text{ almost surely.} \quad (6.16)$$

Suppose for a contradiction that  $X_n^{\geq 3}$  does not go to infinity almost surely. That is, there exists a positive constant  $c$  such that

$$\mathbf{P} \{\forall n \in \mathbb{N}, X_n^{\geq 3} < \infty\} = c > 0.$$

By continuity of probability, this implies that there exists some constant  $K$  such that

$$\mathbf{P} \{\forall n \in \mathbb{N}, X_n^{\geq 3} \leq K\} \geq \frac{c}{2} > 0. \quad (6.17)$$

Using the fact that  $X_n^{\geq 2}$  goes to infinity almost surely and  $X_n^2 = X_n^{\geq 2} - X_n^{\geq 3}$  for all  $n$ , we have

$$\mathbf{P} \{X_n^2 \rightarrow \infty \mid \forall n \in \mathbb{N}, X_n^{\geq 3} \leq K\} = 1. \quad (6.18)$$

Define  $\tau$  to be the smallest time after which, for all  $n \geq \tau$ ,  $T_n$  always contains at least two neighbouring vertices each of degree two,

$$\tau := \inf\{m: \forall n \geq m, \exists u, v \in T_n, u \sim v, d_n(u) = d_n(v) = 2\}.$$

It follows from (6.18) that  $\mathbf{P}\{\tau < \infty \mid \forall n \in \mathbb{N}, X_n^{\geq 3} \leq K\} = 1$ . Note that  $\Delta X_n^{\geq 3} = 1$  if and only if vertex  $n+1$  attaches to a vertex of degree two. At time  $n \geq \tau$  this occurs with probability at least  $1/n$ . Thus, except for finitely many  $n$ ,

$$\mathbf{E}[\Delta X_n^{\geq 3} \mid T_n, X_n^{\geq 3} \leq K] \geq \frac{1}{n};$$

this implies that  $\mathbf{P}\{\lim_n X_n^{\geq 3} > K\} = 1$ , which contradicts our hypothesis.  $\square$

**Lemma 6.18.** As  $n \rightarrow \infty$ ,  $X_n^{\geq 2}/\log n \rightarrow \infty$  almost surely.

*Proof.* Fix  $C > 0$ . We will show that

$$\liminf_{n \rightarrow \infty} \frac{X_n^{\geq 2}}{\log n} \geq C \text{ almost surely,}$$

which implies the statement. Conditionally on  $T_n$ ,  $\Delta X_n^{\geq 2}$  is a Bernoulli random variable with parameter  $\mathbf{E}[\Delta X_n^{\geq 2} \mid T_n]$ . We prove that  $\mathbf{E}[\Delta X_n^{\geq 2} \mid T_n] < C/n$  only finitely many times, in order to couple  $(\Delta X_n^{\geq 2})_{n \geq 1}$  to a sequence of independent Bernoulli random variables with parameter  $C/n$ .

Let  $(U_i, i \geq 1)$  be independent uniform random variables on  $[0, 1]$ . Conditionally on  $T_n$ , construct  $T_{n+1}$  from  $T_n$  by setting  $\Delta X_n^{\geq 2} = 1$  if and only if  $U_n < \mathbf{E}[\Delta X_n^{\geq 2} \mid T_n]$ , and then sampling the additional randomness required to construct  $T_{n+1}$  conditionally on  $T_n$  and on the value of  $\Delta X_n^{\geq 2}$ . Define a coupling between  $(\Delta X_n^{\geq 2})_{n \geq 1}$  and  $(B_n)_{n \geq 1}$ , a sequence of independent Bernoulli random variables with parameter  $C/n$ , by setting  $B_n = 1$  if  $U_n < C/n$  and  $B_n = 0$  otherwise. It is immediate that  $\Delta X_n^{\geq 2} \geq B_n$  whenever  $\mathbf{E}[\Delta X_n^{\geq 2} \mid T_n] \geq C/n$ .

Lemma 6.15 states that  $\mathbf{E}[\Delta X_n^{\geq 2} \mid T_n] \geq \frac{1}{3n} X_n^{\geq 3}$ , and by Lemma 6.17,  $X_n^{\geq 3} \rightarrow \infty$  almost surely, thus

$$\mathbf{E}[\Delta X_n^{\geq 2} \mid T_n] = \omega(1/n) \tag{6.19}$$

almost surely. Therefore, almost surely,  $\mathbf{E}[\Delta X_n^{\geq 2} \mid T_n] < C/n$  only finitely many times, and also  $\Delta X_n^{\geq 2} < B_n$  only finitely many times. In particular

$$\liminf_{n \rightarrow \infty} \frac{X_n^{\geq 2}}{\log n} \geq \liminf_{n \rightarrow \infty} \frac{\sum_{i=1}^n B_i}{\log n}$$

almost surely. We claim that

$$\frac{\sum_{i=1}^n B_i}{\log n} \rightarrow C \text{ almost surely.} \tag{6.20}$$

Indeed, let  $(Y_j)_{j \geq 1}$  be a sequence of independent random variables satisfying  $Y_j = \sum_{i=\lfloor e^{j-1} \rfloor + 1}^{\lfloor e^j \rfloor} B_i$ . Then  $\lim_{j \rightarrow \infty} \mathbf{E}[Y_j] = C$  and  $\mathbf{E}[Y_j^2] \leq 10C^2$ . By Kolmogorov's strong law of large numbers [8, Theorem 3.2.],

$$\frac{\sum_{j=1}^n Y_j}{n} \rightarrow C \text{ almost surely.}$$

This implies the convergence of (6.20) along the subsequence  $(\lfloor e^j \rfloor)_{j \geq 1}$ , and, by monotonicity, (6.20) follows.  $\square$

**Corollary 6.19.** For  $\alpha = (\sqrt{13} - 3)/2$  the unique positive solution of  $x = \frac{1-2x}{1+x}$ , for any  $\beta \in (0, \alpha)$  and for all  $n$  sufficiently large,  $\mathbf{E}[X_n^{\geq 3}] \geq \beta \mathbf{E}[X_n^{\geq 2}]$ .

By Lemma 6.16 we have  $\mathbf{E}X_n^{\geq 3} \geq \beta \mathbf{E}X_n^{\geq 2} - c \log(n)$ , and by Lemma 6.18  $X_n^{\geq 2} = \omega(\log n)$  almost surely; Corollary 6.19 follows.

**Proposition 6.20.** For  $\alpha = (\sqrt{13} - 3)/2$  the unique positive solution of  $x = \frac{1-2x}{1+x}$ , for any  $0 < \delta < \alpha/3 \approx 0.101$  we have  $n^{-\delta} X_n^{\geq 2} \rightarrow \infty$  almost surely.

*Proof.* By Lemmas 6.15 and 6.16, for all  $\beta \in (0, \alpha)$  there exists a  $c > 0$  such that

$$\mathbf{P} \left\{ \mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] < \frac{\beta}{3n} (X_n^{\geq 2} - c \log n) \text{ infinitely often} \right\} = 0. \quad (6.21)$$

Fix  $\delta$  such that  $0 < \delta < \beta/3$  and fix  $\gamma$  rational such that  $\delta < \gamma < \beta/3$ . Lemma 6.18 states that  $X_n^{\geq 2} = \omega(\log n)$  almost surely, therefore, almost surely there are only finitely many  $n$  such that

$$\frac{\beta}{3} (X_n^{\geq 2} - c \log n) < \gamma X_n^{\geq 2},$$

Define the time  $\tau$  as

$$\tau = 3 \vee \sup \left\{ n \geq 1 : \mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] < \frac{\gamma}{n} X_n^{\geq 2} \right\}.$$

As a consequence of (6.21),  $\tau < \infty$  almost surely. For  $n > \tau$ , by the definition of  $\tau$ , the probability that  $X_n^{\geq 2}$  increases at time  $n$  is bounded from below by  $\gamma X_n^{\geq 2}/n$ . It is therefore natural to compare  $X_n^{\geq 2}$  to a generalised Pólya urn. We first introduce an urn process, containing  $B_n$  black balls, and show that  $n^{-\delta} B_n \rightarrow \infty$  almost surely. Conditionally on  $\tau = t$ , we then couple the sequences  $(X_n^{\geq 2})_{n \geq t}$  and  $(B_n)_{n \geq t}$  and conclude the proof of Proposition 6.20.

To introduce the urn process, let  $M > 0$  be an integer such that  $\gamma^{-1}M$  is a positive integer, and let  $t > 0$  be another integer. Consider the urn process started at time  $t$  with  $M$  black balls and  $t\gamma^{-1}M - M$  white balls. At every time step, draw a ball from the urn uniformly at random and return it to the urn together with  $\gamma^{-1}M$  additional balls. If the drawn ball is white, all of the additional balls are white. If the drawn ball is black,  $M$  of the additional balls are black and the other  $(\gamma^{-1} - 1)M$  balls are white. Denote by  $B_n$  the number of black balls in the urn at time  $n$ . Then at time  $n$ , the number of black balls  $B_n$  increases by  $M$  with probability  $B_n/(n\gamma^{-1}M)$ . The described urn is *triangular* since if a white ball is drawn only white balls are added to the urn. The asymptotic behaviour of triangular urn processes has been studied by Janson [12]. Theorem 1.3.(v) in [12] implies that  $n^{-\gamma} B_n$  converges almost surely to some random variable  $Z$ , and Theorem 8.7. in [12] shows that  $Z$  puts no mass on 0, so  $n^{-\delta} B_n \rightarrow \infty$  almost surely as  $\delta < \gamma$ .

We now introduce a coupling satisfying that  $(X_n^{\geq 2})_{n \geq t}$  grows at least as fast as  $(B_n)_{n \geq t}$ , on the event  $\tau < t$ . To formalise this coupling, note that conditionally on  $\tau = t$ , if  $n \geq t$ , then  $\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \geq \frac{\gamma}{n} X_n^{\geq 2}$ . Since  $\Delta X_n^{\geq 2} \in \{0, 1\}$ ,

$$\mathbf{P} \left\{ \Delta(MX_n^{\geq 2}) = M \mid T_n \right\} \geq \frac{MX_n^{\geq 2}}{n\gamma^{-1}M}.$$

Remark that  $MX_t^{\geq 2}$  is at least  $M$ . Let  $(U_n)_{n \geq 1}$  be a sequence of independent uniform random variables on  $[0, 1]$ . Conditionally on  $T_n$ , construct  $T_{n+1}$  from  $T_n$  by setting  $\Delta X_n^{\geq 2} = 1$  if and only if  $U_n \leq \mathbf{E} [\Delta X_n^{\geq 2} \mid T_n]$  and sampling the remaining randomness in  $T_{n+1}$  conditional on  $T_n$  and the value of  $\Delta X_n^{\geq 2}$ . Let  $B_t = 0$  and for  $n \geq t$  let  $\Delta B_n = M$  if  $B_n/(n\gamma^{-1}M) \leq U_n$  and  $\Delta B_n = 0$  otherwise. Then  $(B_n)_{n \geq t}$  is distributed as the number of black balls in the Pólya urn described above. We already noted that  $n^{-\delta} B_n \rightarrow \infty$  almost surely, and by our coupling we have that on the event  $\tau < t$ ,  $MX_n^{\geq 2} \geq B_n$  for all  $n \geq t$ . Thus, for  $\varepsilon > 0$

$$\mathbf{P} \{n^{-\delta} X_n^{\geq 2} \rightarrow \infty\} > \mathbf{P} \{\tau < t\} \geq 1 - \varepsilon.$$

Since  $\tau$  is almost surely finite,  $\varepsilon$  can be chosen arbitrarily small, by taking  $t$  large, which concludes the proof.  $\square$

**Lemma 6.21.** Let  $\gamma = 3 - 2\sqrt{2} \approx 0.172$  be the unique positive solution to  $x = \frac{1-5x}{1-x}$ . Then for any  $\beta \in (1 - \gamma, 1)$  there exists  $C > 0$  such that

$$\mathbf{P} \{X_n^{\geq 3} > \beta X_n^{\geq 2} + C \log n \text{ infinitely often}\} = 0$$

and for all  $n$ ,  $\mathbf{E} [X_n^{\geq 3}] \leq \beta \mathbf{E} [X_n^{\geq 2}] + C \log n$ .

*Proof.* We apply Proposition 6.11 to the sequences  $X_k = X_n^{\geq 2}$  and  $Y_k = X_n^{\geq 3}$ . Note that  $\Delta X_n^{\geq 2} + \Delta X_n^{\geq 3} \in \{0, 1\}$ . It remains to show that there exists  $\gamma > 0$  such that for all  $n$ , either

$$X_n^{\geq 2} \geq \frac{1}{1-\gamma} X_n^{\geq 3} \text{ or}$$

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \geq \frac{1}{1-\gamma} \mathbf{E} [\Delta X_n^{\geq 3} \mid T_n].$$

Lemma 6.21 follows then directly by applying Proposition 6.11. Suppose that  $X_n^{\geq 2} < \frac{1}{1-\gamma} X_n^{\geq 3}$ . Since  $X_n^2 = X_n^{\geq 2} - X_n^{\geq 3}$ , (6.5) gives that

$$\mathbf{E} [\Delta X_n^{\geq 3} \mid T_n] \leq \frac{4}{3n} (X_n^{\geq 2} - X_n^{\geq 3}) < \frac{4\gamma}{3n} X_n^{\geq 2}.$$

From Lemma 6.15 we see that

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \geq \frac{1}{3n} X_n^{\geq 3} > \frac{1-\gamma}{3n} X_n^{\geq 2}.$$

Therefore,

$$\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] > \frac{1-\gamma}{4\gamma} \mathbf{E} [\Delta X_n^{\geq 3} \mid T_n].$$

The statement follows from the choice of  $\gamma$ .  $\square$

**Proposition 6.22.** Let  $\delta = 1 - \gamma/2$ , where  $\gamma = 3 - 2\sqrt{2}$ . Then, for any  $\delta < \lambda < 1$ ,

$$n^{-\lambda} X_n^{\geq 2} \rightarrow 0,$$

almost surely.

*Proof.* By (6.4),  $\mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] \leq \frac{1}{2n} X_n^{\geq 2} + \frac{1}{2n} X_n^{\geq 3}$ . Combining this with Lemma 6.21 gives that for any  $\beta \in (1 - \gamma, 1)$ , there exists a  $C > 0$  such that

$$\mathbf{P} \left\{ \mathbf{E} [\Delta X_n^{\geq 2} \mid T_n] > \frac{1+\beta}{2n} X_n^{\geq 2} + C \log n \text{ infinitely often} \right\} = 0.$$

By mimicking the proof of Proposition 6.20, we can compare  $X_n^{\geq 2}$  to a generalised Pólya urn and obtain an upper bound on  $\Delta X_n^{\geq 2}$ . Omitting the details of the coupling, we conclude that for  $\lambda \in (\frac{1+\beta}{2}, 1)$  it holds that  $\mathbf{P} \{n^{-\lambda} X_n^{\geq 2} \rightarrow 0\} = 1$ .  $\square$

*Proof of Theorem 4.7 and 4.8.* The second part of Theorem 4.7 follows from Proposition 6.22, by noting that  $X_n^{\geq k} \leq X_n^{\geq 2}$  and therefore, for any  $k \geq 2$ ,

$$n^{-\lambda} X_n^{\geq k} \rightarrow 0 \text{ a.s.}$$

The upper bound in Theorem 4.8 follows directly since  $X_n^{\geq k+1} \leq X_n^{\geq k}$  for all  $k$ .

By Proposition 6.20,  $\lim_n n^{-\delta} X_n^{\geq 2} = \infty$  almost surely. We prove the remaining cases in the first part of Theorem 4.7 and the lower bound of Theorem 4.8, by using induction to prove that, almost surely, for all  $k \geq 2$ ,

- (i) there exists a positive constant  $c_k$  such that  $\liminf_n X_n^{\geq k+1}/X_n^{\geq k} > c_k$ , and
- (ii)  $\lim_n n^{-\delta} X_n^{\geq k+1} = \infty$ .

The fact that  $\lim_n n^{-\delta} X_n^{\geq 2} = \infty$  almost surely and Lemma 6.16 imply that (i) holds for  $k = 2$  for any  $0 < c_2 < \alpha$ . This then also implies (ii) for  $k = 2$ .

Now, fix  $k \geq 3$  and suppose that the induction hypothesis holds for all  $2 \leq \ell \leq k - 1$ . Let  $b_k \in (0, 1)$  be the solution to

$$\frac{b_k}{1 - 2b_k} = \frac{c_{k-1}(k-1)}{k(k-3/2)},$$

and fix  $0 < c_k < b_k$ . We claim that for all  $n$ , either

$$X_n^{\geq k+1} \geq b_k X_n^{\geq k} \text{ or} \\ \mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] \geq b_k \mathbf{E} [\Delta X_n^{\geq k} \mid T_n]$$

almost surely, except at finitely many times. If this holds, then applying Proposition 6.11 gives us that there exists  $C > 0$  such that

$$\mathbf{P} \{X_n^{\geq k+1} < c_k X_n^{\geq k} - C \log n \text{ infinitely often}\} = 0.$$

By the induction hypothesis,  $n^{-\delta} X_n^{\geq k} \rightarrow \infty$  and so

$$\mathbf{P} \{X_n^{\geq k+1}/X_n^{\geq k} < c_k \text{ infinitely often}\} = 0.$$

This implies that (i) holds at step  $k$ , which in turn implies part (ii). It remains to prove the claim. Suppose that

$$X_n^{\geq k+1} < b_k X_n^{\geq k}.$$

From (6.7),  $b_k X_n^{\geq k} > X_n^{k, > k}$ , which combined with (6.6) implies that

$$X_n^k < b_k X_n^{\geq k} + X_n^{k, \leq k}.$$

Now, using (6.8) gives

$$\mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] > \frac{k-1}{kn} (X_n^k - b_k X_n^{\geq k}) = \frac{k-1}{kn} ((1-b_k)X_n^{\geq k} - X_n^{\geq k+1}),$$

where the equality holds since  $X_n^k = X_n^{\geq k} - X_n^{\geq k+1}$ . Our assumption  $X_n^{\geq k+1} < b_k X_n^{\geq k}$  then gives

$$\mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] \geq \frac{k-1}{kn} (1 - 2b_k) X_n^{\geq k}.$$



The induction hypothesis for  $k - 1$  implies that almost surely  $X_n^{\geq k} > c_{k-1} X_n^{\geq k-1} \geq c_{k-1} X_n^{k-1}$ , except at finitely many times. Therefore,

$$\mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] \geq \frac{c_{k-1}(k-1)}{kn} (1 - 2b_k) X_n^{k-1}$$

except at finitely many times. But (6.9) implies that for all  $n$  sufficiently large we have that  $X_n^{k-1} \geq \frac{k-3/2}{n} \mathbf{E} [\Delta X_n^{\geq k} \mid T_n]$ , so we conclude that, except at finitely many times

$$\mathbf{E} [\Delta X_n^{\geq k+1} \mid T_n] \geq \frac{c_{k-1}(k-1)}{k(k-3/2)} (1 - 2b_k) \mathbf{E} [\Delta X_n^{\geq k} \mid T_n] .$$

This proves the claim by our choice of  $b_k$ , which concludes the proof.  $\square$

## 7 Open questions and future directions

We conclude with some open questions about the random friend tree.

1. In Theorem 4.7 we prove that, for some  $0.1 < \delta < \lambda < 0.9$ , almost surely  $n^\delta \ll X_n^{\geq 2} \ll n^\lambda$ . A question of interest would be whether the gap between the upper and lower bound can be closed and whether, for some  $\mu > 0$  and some random variable  $X$  with non-trivial support,  $n^{-\mu} X_n^{\geq 2} \rightarrow X$  almost surely. Simulations by Krapivsky and Redner [15] suggest that  $X_n^{\geq 2}$  grows as  $n^\mu$ , with  $\mu \approx 0.566$ .
2. We prove that, for fixed  $k$ , the number of vertices with degree at least  $k + 1$  is of the same order as the number of vertices with degree at least  $k$ , see Theorem 4.8. Can we prove, for fixed  $k$ , that the number of degree- $k$  nodes is of the same order as the number of degree- $(k + 1)$  nodes? Does it hold that  $\limsup_{n \rightarrow \infty} \frac{X_n^{\geq k}}{X_n^{\geq k+1}}$  goes to 0 as  $k$  goes to infinity? Or, informally, are most of the non-leaves vertices of bounded degree? Krapivsky and Redner [15] conjecture that for each  $k$ ,  $\frac{X_n^k}{X_n^{\geq 2}}$  has an almost sure limit that is  $\Theta(k^{-(1+\mu)})$ .
3. Is  $\mathbf{P} \{Z_u > 0\}$  decreasing in  $u$ ? More generally, does  $Z_u$  stochastically dominate  $Z_v$  for  $u < v$ ?
4. Does it hold that  $\sum_{i \geq 1} Z_i = 1$  almost surely?
5. We know that every edge contains a vertex of linear degree, but the diameter of the tree grows logarithmically, so there must be connected subtrees consisting of low-degree vertices whose linear growth has not kicked in yet. This is illustrated by the proof of Proposition 6.8, which shows that there are paths of length  $\Theta(\log(n)/\log \log(n))$  that consist of just degree 2 vertices. It would be interesting to get a better understanding of the law of these exceptional substructures that contain most of the low-degree vertices. What does the forest induced by the vertices of degree at most  $N$ , for large  $N$ , look like? Do these subtrees look like ‘young’ friend trees?
6. A natural extension of the model is to attach the new vertex to multiple, say  $m$ , vertices. There are two variants: either  $V_n$  is a uniformly random vertex and the new vertex  $n + 1$  attaches to  $m$  independently sampled random neighbours of  $V_n$ , or we let  $V_n^{(1)}, \dots, V_n^{(m)}$  to be independent random vertices and we let  $n + 1$  connect to a uniform neighbour of each of the  $V_n^{(i)}$ .
7. A second variation is to choose  $0 < p < 1$  and connect to  $V_n$  with probability  $p$  and to  $W_n$  with probability  $1 - p$ . This modification makes it much easier for neighbours of high-degree vertices to grow their degree, and in particular, the degree of every

vertex goes to infinity almost surely as the tree grows. It would be interesting to see how much of the structure of the random friend tree remains after this modification.

8. Another final modification of the model, as described in the introduction, is to let  $W_n$  be the endpoint of a random walk with  $k$  steps rather than 1 step from  $V_n$ . In the case of  $k = 0$ , we obtain an URRT and if  $k$  is sufficiently large such that the random walk is perfectly mixed, we get a PA tree. One could study how properties such as the size of the largest degree depend on  $k$ .

## 8 Appendix

We prove Proposition 6.11. We start by stating and proving a technical lemma that is needed for its proof.

We make use of the following straightforward fact. Fix  $a, b > 0$  and let  $(Y_k)_{k \geq 0}$  be a random walk with steps in  $\{-a, b\}$  such that  $\mathbf{E}[\Delta Y_k] = c > 0$ . Then,  $\mathbf{P}\{\Delta Y_k = b\} = (a + c)/(a + b)$  and by writing  $\tau = \inf\{k \geq 0 : Y_k < 0\}$ , we have  $\mathbf{P}\{\tau = \infty\} > 0$ .

**Lemma 8.1.** Let  $B = (B_k)_{k \geq 0}$  be a random process adapted to a filtration  $(\mathcal{F}_k)_{k \geq 0}$ . Suppose that there exist  $a, b > 0$  such that, almost surely for each  $k$ ,  $\Delta B_k \in \{-a, 0, b\}$ . Suppose further that there exists a constant  $c > 0$  such that

$$\mathbf{E}[\Delta B_k \mid \mathcal{F}_k, B_k < 0, \Delta B_k \neq 0] \geq c.$$

Then, there exists a constant  $C = C(a, b, c)$  such that

$$\mathbf{P}\{B_n < -C \log n \text{ infinitely often}\} = 0,$$

and  $\mathbf{E}[B_n] \geq -C \log n$ .

*Proof.* We bound  $B$  from below by another, simpler process  $S = (S_n)_{n \geq 0}$ . The conditions of the lemma imply that we may couple  $B$  with a sequence  $(Y_k)_{k \geq 0}$  of independent, random variables taking values in  $\{-a, b\}$  with

$$\mathbf{P}\{Y_k = b\} = \frac{a + c}{a + b} = 1 - \mathbf{P}\{Y_k = -a\},$$

such that for all  $k \geq 0$ , on the event that  $B_k < 0$  and  $\Delta B_k \neq 0$  we have  $\Delta B_k \geq Y_k$ . We define  $S$  via the following coupling with  $B$ :

1. if  $B_n \geq 0$ , then  $S_{n+1} = -a$ ,
2. if  $B_n < 0$  and  $\Delta B_n = 0$  then  $\Delta S_n = 0$ ,
3. if  $B_n < 0$  and  $\Delta B_n \neq 0$  then  $\Delta S_n = Y_n$ .

An illustration of the coupling can be found in Figure 8. Since  $B_n \geq 0$  implies  $B_{n+1} \geq -a$ , it is immediate that  $S_n \leq B_n$  for all  $n \geq 0$ . The process  $S$  is a sequence of independent negative (incomplete) excursions of a random process that, restricted to the non-constant steps, has independent increments with positive mean except for finitely many times almost surely. There are at most  $n/2$  such excursions by time  $n$ , and if we collapse the constant steps, they are independent realisations of a random walk with step size in  $\{-a, b\}$  and drift  $c$ , started at  $-a$  and ended before (or when) reaching 0. To understand their minimum, let us denote by  $R = (R_n)_{n \geq 0}$  a random walk starting at 0, with steps in  $\{-a, b\}$  and strictly positive drift  $c$ . We define

$$\tau_1 = \inf\{k : R_k < 0\},$$

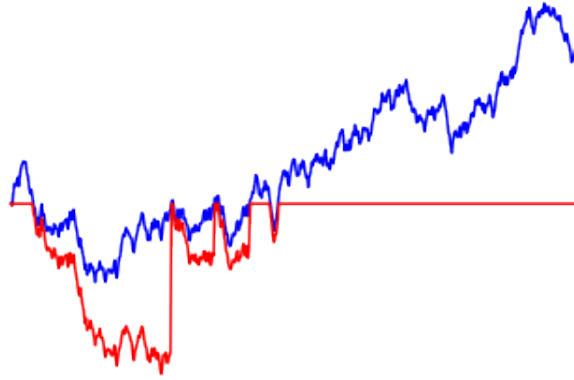


Figure 4: Illustration of a coupling between  $B$  (in blue) and  $S$  (in red).

and

$$\tau_\ell = \inf \{k > \tau_{\ell-1} : R_k < R_{\tau_{\ell-1}}\}.$$

With positive probability,  $R$  stays positive forever, and in particular, using the fact stated just before Lemma 8.1, there exists  $0 < p < 1$  such that

$$0 < 1 - p = \mathbf{P} \{ \tau_\ell = \infty \mid \tau_{\ell-1} < \infty \}.$$

Since the increments are bounded from below by  $-a$ , we know that  $R_{\tau_\ell} - R_{\tau_{\ell-1}} \geq -a$ , which together with the previous identity implies that

$$\min_{k \geq 1} \{R_k\} \succeq_{st} -a \cdot \text{Geom}(1 - p).$$

From the definition of the coupling, we know that  $B_n \geq S_n$  and that  $S_n + a$  stochastically dominates the minimum of  $n$  realisations of  $(R_k)_{k \geq 0}$ . Let  $(A_n)_{n \geq 1}$  be independent  $\text{Geom}(1 - p)$  random variables. Then

$$\mathbf{P} \{S_n \leq -a(k + 1)\} \leq \mathbf{P} \left\{ \max_{i \in [n]} A_i \geq k \right\} \leq np^k.$$

The upper bound is at most  $n^{-2}$  if  $k \geq -3 \log n / \log p$  and the first assertion follows from the Borel–Cantelli lemma. Taking  $k = -3 \log n / \log p + \ell$ , the above bound likewise implies that

$$\mathbf{P} \left\{ -\frac{S_n}{a} + 3 \log n / \log p \geq \ell + 1 \right\} \leq np^{-3 \log n / \log p} p^\ell,$$

and summing over  $\ell \geq 0$  gives the bound

$$\mathbf{E} \left[ -\frac{S_n}{a} + 3 \log n / \log p \right] \leq np^{-3 \log n / \log p} \frac{1}{1 - p},$$

which establishes the second assertion of the lemma.  $\square$

We are now ready to prove the proposition.

*Proof of Proposition 6.11.* We start by proving Proposition 6.11 in the case where the hypothesis holds for all  $n$  (and not all but finitely many  $n$ ). Let  $\beta \in (0, \alpha)$ . We prove the proposition by applying Lemma 8.1 to the process  $B = (B_k)_{k \geq 0}$  with  $B_k := X_k - \beta Y_k$ .

Note that  $B$  has increments in  $\{-\beta, 0, 1\}$ . We need to show that there exists a constant  $c$  such that for all  $k$ ,

$$\mathbf{E}[\Delta B_k \mid \mathcal{F}_k, B_k < 0, \Delta B_k \neq 0] \geq c. \quad (8.1)$$

Define

$$p_k^+ = \mathbf{P}\{\Delta X_k = 1 \mid \mathcal{F}_k, (\Delta X_k, \Delta Y_k) \neq (0, 0)\} = \mathbf{P}\{\Delta B_k = 1 \mid \mathcal{F}_k, \Delta B_k \neq 0\}$$

and so

$$1 - p_k^+ = \mathbf{P}\{\Delta Y_k = 1 \mid \mathcal{F}_k, (\Delta X_k, \Delta Y_k) \neq (0, 0)\} = \mathbf{P}\{\Delta B_k = -\beta \mid \mathcal{F}_k, \Delta B_k \neq 0\}.$$

Note that, if  $\{B_k < 0\}$ , then  $\{X_k < \beta Y_k\}$ , and so  $\mathbf{E}[\Delta X_k \mid \mathcal{F}_k] \geq \alpha \mathbf{E}[\Delta Y_k \mid \mathcal{F}_k]$ , which implies that  $p_k^+ \geq \alpha(1 - p_k^+)$ . We split the event  $\{B_k < 0\}$  into two cases and show that  $\mathbf{E}[\Delta B_k \mid \mathcal{F}_k, \Delta B_k \neq 0]$  is bounded below by some positive constant in both cases. If  $B_k < 0$  and  $p_k^+ \leq \frac{1/2 + \beta}{1 + \beta}$ , then

$$\mathbf{E}[\Delta B_k \mid \mathcal{F}_k, \Delta B_k \neq 0] = p_k^+ - \beta(1 - p_k^+) \geq (\alpha - \beta)(1 - p_k^+) \geq \frac{(\alpha - \beta)}{2 + 2\beta}.$$

On the other hand, if  $B_k < 0$  and  $p_k^+ > \frac{1/2 + \beta}{1 + \beta}$ , then

$$\mathbf{E}[\Delta B_k \mid \mathcal{F}_k, \Delta B_k \neq 0] = p_k^+(1 + \beta) - \beta \geq 1/2.$$

Therefore, (8.1) holds, with a lower bound of  $\min\left\{\frac{(\alpha - \beta)}{2 + 2\beta}, 1/2\right\}$  for  $c$ , and the claim then directly follows from Lemma 8.1.

Finally, we prove that the first statement in the proposition still holds when the assumptions fail at a finite number of times.

We call a time  $k$  *bad* when  $\mathbf{E}[\Delta X_k \mid \mathcal{F}_k] < \alpha \mathbf{E}[\Delta Y_k \mid \mathcal{F}_k]$  and  $\{X_k < \alpha Y_k\}$ ; otherwise we call it *good*. We couple  $(X, Y)$  to a slightly modified process  $(X', Y')$  that has the same increments as  $(X, Y)$  except at bad times, and that satisfies the assumptions at all times. Observe that for each  $k$ , given  $\mathcal{F}_k$ , we know whether  $k$  is bad or not. If  $k$  is bad, set  $(\Delta X'_k, \Delta Y'_k) = (1, 0)$ . If  $k$  is good, set  $(\Delta X'_k, \Delta Y'_k) = (\Delta X_k, \Delta Y_k)$ . We claim that  $(X', Y')$  satisfies the assumptions at all times. The requirement  $0 \leq \Delta X'_k + \Delta Y'_k \leq 1$  for all  $k$  is obviously satisfied. Moreover, for bad  $k$ ,  $\mathbf{E}[\Delta X'_k \mid \mathcal{F}_k] = 1$  and  $\mathbf{E}[\Delta Y'_k \mid \mathcal{F}_k] = 0$ , so at bad times the second requirement is also satisfied. Finally, by construction,  $X'_k \geq X_k$  and  $Y'_k \leq Y_k$  for all  $k$ , so if  $k$  is good and  $\{X'_k < \alpha Y'_k\}$ , then also  $\{X_k < \alpha Y_k\}$ , and therefore

$$\mathbf{E}[\Delta X'_k \mid \mathcal{F}_k] = \mathbf{E}[\Delta X_k \mid \mathcal{F}_k] \geq \alpha \mathbf{E}[\Delta Y_k \mid \mathcal{F}_k] = \alpha \mathbf{E}[\Delta Y'_k \mid \mathcal{F}_k].$$

Then, the first part of the proof implies that for any  $\beta \in (0, \alpha)$  there exists  $C > 0$  such that  $\mathbf{P}\{X'_n < \beta Y'_n - C \log n \text{ i.o.}\} = 0$ . Finally, for  $B$  the total number of bad times, for each  $k$  it holds that  $X'_k \leq X_k + B$  and  $Y'_k \geq Y_k - B$ . This implies that if  $X_k < \beta Y_k - 2C \log k$  then either  $X'_k < \beta Y'_k - C \log n$  or  $2B > C \log k$ . Therefore,

$$\mathbf{P}\{X_n < \beta Y_n - 2C \log n \text{ i.o.}\} \leq \mathbf{P}\{X'_n < \beta Y'_n - C \log n \text{ i.o.}\} + \mathbf{P}\{B = \infty\} = 0. \quad \square$$

## References

- [1] Louigi Addario-Berry and Kevin Ford, *Poisson-Dirichlet branching random walks*, The Annals of Applied Probability **23** (2013), no. 1, 283 – 307.
- [2] Sayan Banerjee, Shankar Bhamidi, and Xiangying Huang, *Co-evolving dynamic networks*, arXiv preprint arXiv:2203.11877 (2022).
- [3] A. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512.

- [4] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities. A nonasymptotic theory of independence*, Oxford University Press, 2013.
- [5] Chris Cannings and Jonathan Jordan, *Random walk attachment graphs*, *Electronic Communications in Probability* **18** (2013), no. none, 1 – 5.
- [6] Luc Devroye, *Applications of the theory of records in the study of random trees*, *Acta Informatica* **26** (1988), no. 1, 123–130.
- [7] Luc Devroye and Jiang Lu, *The strong convergence of maximal degrees in uniform random recursive trees and dags*, *Random Structures & Algorithms* **7** (1995), no. 1, 1–14.
- [8] Richard Durrett, *Essentials of stochastic processes*, 3rd edition 2016. ed., Springer Texts in Statistics, Springer International Publishing, Cham, 2016 (eng).
- [9] Janos Engländer, Giulio Iacobelli, Gábor Pete, and Rodrigo Ribeiro, *Structural results for the tree builder random walk*, arXiv preprint arXiv:2311.18734 (2023).
- [10] Tim S. Evans and Jari Saramäki, *Scale-free networks from self-organization*, *Physical Review E* **72** (2005), 026138.
- [11] Svante Janson, *Asymptotic degree distribution in random recursive trees*, *Random Structures & Algorithms* **26** (2005), no. 1-2, 69–83.
- [12] ———, *Limit theorems for triangular urn schemes*, *Probability Theory and Related Fields* **134** (2006), 417–452.
- [13] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MRMR1782847
- [14] Jon M. Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S. Tomkins, *The web as a graph: Measurements, models, and methods*, Computing and Combinatorics, Springer-Verlag, 1999, pp. 1–17.
- [15] Pavel Krapivsky and Sidney Redner, *Emergent network modularity*, *Journal of Statistical Mechanics: Theory and Experiment* (2017), no. 7, 073405, 22. MR3683821
- [16] Sanjay Kumar, Neeraj Bhat, and BS Panda, *Analysis of social network metrics based on the model of random recursive tree*, *Journal of Interdisciplinary Mathematics* **23** (2020), no. 1, 237–246.
- [17] Hosam Mahmoud, *Pólya urn models*, Chapman and Hall/CRC, 2008.
- [18] Boris Pittel, *Note on the heights of random recursive trees and random m-ary search trees*, *Random Structures & Algorithms* **5** (1994), no. 2, 337–347.
- [19] Jari Saramäki and Kimmo Kaski, *Scale-free networks generated by random walkers*, *Physica A: Statistical Mechanics and its Applications* **341** (2004), 80–86.
- [20] Remco van der Hofstad, *Random graphs and complex networks*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2016.
- [21] ———, *Random graphs and complex networks: Volume 2*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2024.

**Acknowledgments.** Louigi Addario-Berry is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chairs program.

Simon Briend is supported by Région Île de France.

Luc Devroye is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Serte Donderwinkel is supported by the CogniGron research center and the Ubbo Emmius Funds (Univ. of Groningen).

Céline Kerriou is supported by the DFG project 444092244 “Condensation in random geometric graphs” within the priority programme SPP 2265.

Gábor Lugosi is supported by Ayudas Fundación BBVA a Proyectos de Investigación Científica 2021 and the Spanish Ministry of Economy and Competitiveness grant PID2022-138268NB-I00, financed by MCIN/AEI/10.13039/501100011033, FSE+MTM2015-67304-P, and FEDER, EU.